THE UBIQUITY OF IODOTYROSINE DEIODINASE: IDENTIFICATION OF ITS SIGNATURE SEQUENCE AND FUNCTIONAL ANALYSIS IN A MODEL INVERTEBRATE

by
Abhishek V Phatarphekar

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

August, 2016

Abstract

Iodotyrosine deiodinase (IYD) serves a critical role in iodide conservation through salvage of iodide from mono- and diiodotyrosine (I-Tyr and I_2-Tyr), the by-products of thyroid hormone (TH) synthesis. Genomic data suggested the presence of this enzyme in a wide range of organisms including non-chordate invertebrates as well as prokaryotes. These organisms are not associated with TH signaling. A representative set of IYD homologs from these organisms with conserved key residues were expressed, purified and shown to catalyze deiodination of I_2-Tyr with comparable catalytic efficiencies (k_{cat}/K_m). Our findings implied that the origins of IYD predate the evolution of TH signaling and suggested alternative physiological roles for IYD in non-chordate invertebrates.

The IYD homolog from the model non-chordate invertebrate *Drosophila melanogaster* (dmIYD) displayed little discrimination between iodo-, bromo- and chlorotyrosine as substrates in vitro since all exhibited similar affinities and comparable efficiencies of dehalogenation. Additionally, active site mutants of dmIYD revealed a glutamate to be most critical for catalysis followed by a lysine while substitution of a tyrosine was least perturbing. These findings should further facilitate annotation of genomic data for potential halotyrosine dehalogenases.

The in vivo function of the dehalogenase in non-chordate invertebrates is unknown. Prior expression profiling of mRNA in *Drosophila* indicated that testes expressed the highest level of the dehalogenase encoding gene (CG6279). This was confirmed through in situ hybridization experiments suggesting a biological role for CG6279 in spermatogenesis. However, deletion of the CG6279 gene ($CG6279^{Del}$) as well as

introduction of a loss of function amino acid substitution in the dehalogenase domain (*CG6279*^{E154Q}) using CRISPR/Cas9 gene editing did not significantly affect the fertility of male *Drosophila.* This indicated that the dehalogenase gene is not critical for spermatogenesis.

Previous studies demonstrated I-Tyr to be toxic to *Drosophila* larvae. To assess whether loss of function mutations in CG6279 increases larval sensitivity to I-Tyr toxicity, 1st instar larvae from both CG6279 mutant strains were raised on media supplemented with I-Tyr and survival to adulthood was evaluated. *CG6279*^{E154Q} larvae displayed significantly lower survival rates in presence of I-Tyr compared to controls indicating increased sensitivity to I-Tyr toxicity. However, *CG6279*^{Del} larvae were significantly less sensitive to I-Tyr toxicity leading to uncertainty whether the dehalogenase is metabolically active in vivo.

Advisor: Prof. Steven E. Rokita

Readers: Prof. Craig A. Townsend

Prof. Marc M. Greenberg

Dedication

To my family, for their unconditional support

Acknowledgments

My scientific career has been influenced by many people. None have had as profound an impact as by advisor and mentor Professor Steven E. Rokita. Firstly, I would like to thank him for accepting me into his research group while we were still located at University of Maryland and enabling me to move with him to Johns Hopkins University. I appreciate his efforts towards training me to think like a scientist especially regarding those crucial control experiments, for sound advice and also his incessant encouragement. I would like to specially thank him for allowing me to steer this project into the 'unknown realm' of biology relatively unfamiliar to us biochemists. Lastly, but importantly thank you for allowing me to maintain a healthy work-life balance over the past 5 years. I would like to thank Professor Craig Townsend and Professor Marc Geenberg for serving on my thesis advisory committee and taking the time to read and evaluate my work. A special mention to my former manager and friend at Merck, Dr Robert Davidson for being supportive and encouraging of my decision to return to graduate school to complete my PhD.

My work would not have been possible without the guidance, assistance and facilities of our collaborators, Professor Xin Chen and her group at the Biology Dept. I would like to thank Dr Chen and her former postdoc Dr Sukho Eun for practically teaching me *Drosophila* genetics and training me in the required lab skills. I would also like to thank Lijuan Feng for helping me troubleshoot various fly issues. I would also like to thank other members of the Chen lab for accepting me into their fold. In addition, I would like to thank Dr Shekerah Primus from the Van Doren lab and Professor Robert Johnston and

his graduate student Caitlin Anderson for help with design of RNAi and CRISPR experiments respectively.

I would like to thank all the present as well as the former members of the Rokita lab. I am especially grateful to Nattha Ingavat, Blessing Deeyaa and Mark Hutchinson. All four of us joined the lab together in the same year and have been through the ups and downs together. Thank you for being good lab-mates and friends. I would like to thank Dr Shalini Saha for sharing her molecular biology and protein chemistry knowledge whenever needed and for great conversations in general even regarding my work with *Drosophila*. Thank you to Dr Jen Buss, Dr Petrina Boucher and Dr Fazel Fakhari for sharing advice as senior graduate students when I joined the lab. I would also like to acknowledge Dr Arnab Mukherjee for good times outside of the lab. To all the young members of the Rokita lab, thank you for taking over our responsibilities.

I am eternally grateful towards my family and my friends for all their love, support, encouragement and contributions. To my grandparents who have showered me with love. To my wife for patiently listening to me raving and ranting about science and unconditionally supporting me. To my parents for supporting me in all my endeavors through all the means available. To my lovely daughter whose arrival added tremendous joy and her smile made all those failed experiments so much more forgettable.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Chapter 2: Refining the signature sequence for IYD activity through heterologous expression and characterization of IYD homologs

Chapter 3: Functional analysis of the *Drosophila* halotyrosine dehalogenase

Chapter 4: Silencing the halotyrosine dehalogenase gene in *Drosophila* using RNAi

Chapter 5: CRISPR/Cas9-based editing of a *Drosophila* dehalogenase gene

Chapter 6: Probing the in vivo activity of the *Drosophila* halotyrosine dehalogenase

List of Abbreviations

AP – Alkaline phosphatase

BDSC - Bloomington *Drosophila* stock center

BLAST - Basic local alignment search tool

bp - Base pairs

Br-Tyr - 3-bromo-L-tyrosine

CG6279 - Gene encoding halotyrosine dehalogenase in *Drosophila*

$CG6279^{Del}$ – *Drosophila* strain with deletion of the dehalogenase gene

$CG6279^{E154Q}$ - *Drosophila* strain with a E154Q codon mutation in the dehalogenase gene

$CG6279^{WT}$ - *Drosophila* strain with the wild-type dehalogenase gene

CG6279-RA - Transcript isoform A from gene CG6279

CG6279-RB - Transcript isoform B from CG6279

Cl-Tyr - 3-chloro-L-tyrosine

CRISPR - Clustered regularly interspaced palindromic repeats

Df - Deficiency genetic background for gene CG6279 in *Drosophila*

dmIYD - IYD from *Drosophila melanogaster*

dpIYD - IYD from *Daphnia pulex* (daphnia)

DSB - Double stranded breaks

DTT - Dithiothreitol

EDTA - Ethylenediaminetetraacetic acid

FMN - Flavin mononucleotide

FRP - Flavin reductase P

F-Tyr - 3-fluoro-L-tyrosine

gRNA - Guide RNA for CRISPR editing

HDR - Homology directed repair

hmIYD - IYD from *Hydra magnipapillata* (hydra)

hhIYD - IYD from the *Haliscomenobacter hydrossis* DSM 1100

HPLC - High performance liquid chromatohraphy

hsIYD - IYD homolog from *Homo sapien* (human)

ID - Iodothyronine deiodinase

IPTG - Isopropyl-β-D-thiogalacto-pyranoside

I-Tyr - 3-Iodo-L-tyrosine

I_2-Tyr - 3, 5-Diiodo-L-tyrosine

IYD - Iodotyrosine deiodinase

L1 – 1st instar of larval stage of *Drosophila*

mmIYD - IYD from *Mus musculus* (mouse)

MWCO - Molecular weight cut off

NCBI - National Center for Biotechnology Information

Ni-NTA - Nitrilotriacetic acid

NIS - Sodium / iodide symporter

NOX - NADH oxidase

Nt - Nucleotides

nvIYD - IYD from *Nematostella vectensis* (sea anemone)

PAM - Protospacer adjacent motif

PCR - Polymerase chain reaction

pfIYD - IYD from *Pyrococcus furiosus*

SEC - Size exclusion chromatography

shmiRNA - Short hairpin micro RNA

siRNA - Small interfering RNA

ssODN - Single strand oligodeoxynucleotide

stIYD - IYD from *Streptomyces Sp HPH 0547*

SUMO - Small ubiquitin-related modifier

TCEP - Tris (2-carboxyethyl)phosphine

THs - Thyroid hormones

tnIYD - IYD from *Thermotoga neapolitana*

TPO - Thyroid peroxidase

UAS - Upstream activating sequence

UTR - Untranslated region

VDRC - Vienna *Drosophila* RNAi Center

vhIYD - IYD from *Vibrio harveyi CAIM 1792*

WT - Wild-type

yw - Yellow-White *Drosophila* melanogaster strain

λ_{ex} - Excitation wavelength

λ_{em} - Emission wavelength

Chapter 1: Introduction

1.1 Iodine is an essential micronutrient

Iodine is critical for human health. A lack of iodine may lead to goiter (enlargement of the thyroid) and thyroid hormone (TH) deficiency which can result in serious developmental disorders in early childhood.[1,2] Intake of iodine recommended by the World Health Organization for humans 12 years and above is about 150 µg per day.[2] Despite this relatively small requirement, iodine can be hard to come by since it is the least abundant halogen in the earth's crust and its concentration is especially low in mountainous and inland areas.[1,3] Even the concentration of iodine in the oceans (as iodide or I^-) is quite low and estimated to be about 50 µg/L.[1,4]

Iodine is essential for the synthesis of THs namely 3,3',5,5'-tetraiodothyronine (T4) and 3,3',5-triiodothyronine (T3) (Fig. 1-1).[2] THs are critical for vertebrate development. In mammals, they regulate growth, development, differentiation of nervous tissue and the basal metabolic rate.[5] In non-mammalian vertebrates such as amphibians, THs play a critical role in regulating metamorphosis.[6] Removal of the thyroid in amphibian larvae has been shown to prevent metamorphosis.[7] In bony fish like the zebrafish, THs control development from the early embryonic stages.[8] Even chordate invertebrates such as the amphioxus (lancelet) have an endostyle (an evolutionary precursor of thyroid) and likely produce THs that again can trigger metamorphosis.[9,10]

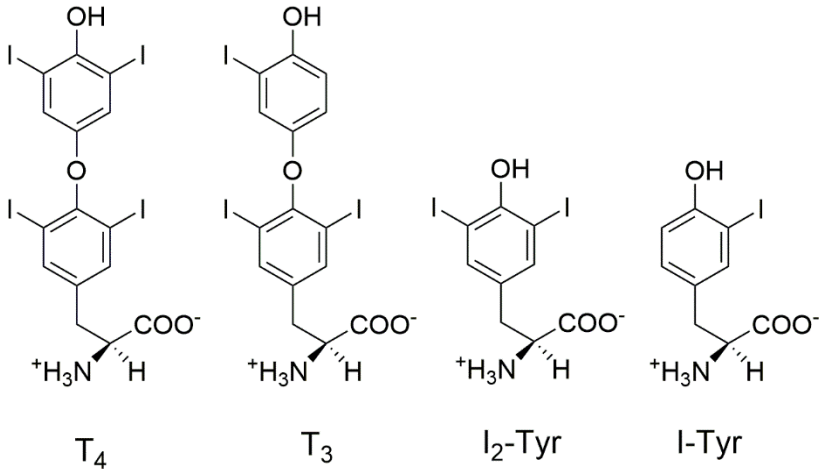

Figure 1-1. Thyroid hormones T4 and T3 along with diiodotyrosine (I_2-Tyr) and monoiodotyrosine (I-Tyr).

1.2 Iodide uptake and conservation

Mammals have evolved to efficiently conserve and recycle iodide (Fig. 1-2). It is concentrated in thyroid follicular cells through active transport by the sodium/iodide symporter (NIS), a basal membrane bound glycoprotein.[11] This results in a 20 to 40-fold higher concentration of I^- within the thyroid follicular cells compared to plasma.[12] I^- is then transported to the colloid through the apical membrane where it is coupled to tyrosine residues on thyroglobulin in the colloid by thyroid peroxidase (TPO). TPO also couples adjacent iodinated tyrosines to produce TH precursors.[13] The thyroglobulin then undergoes endocytosis into the follicular cell where it is degraded to release one molecule of TH along with an estimated 7 molecules of monoiodotyrosine (I-Tyr) and about 6 molecules of diiodotyrosine (I_2-Tyr) that are iodinated by-products of TH synthesis (Fig 1-1).[13] I-Tyr and I_2-Tyr cannot be directly reused for TH synthesis but iodotyrosine deiodinase (IYD) can salvage I^- from these by-products of TH synthesis.[14] IYD is essential for maintaining adequate levels of iodide for TH synthesis and mutations in this gene have been shown to

result in congenital hypothyroidism with severe goiter and even cognitive defects in humans.[15]

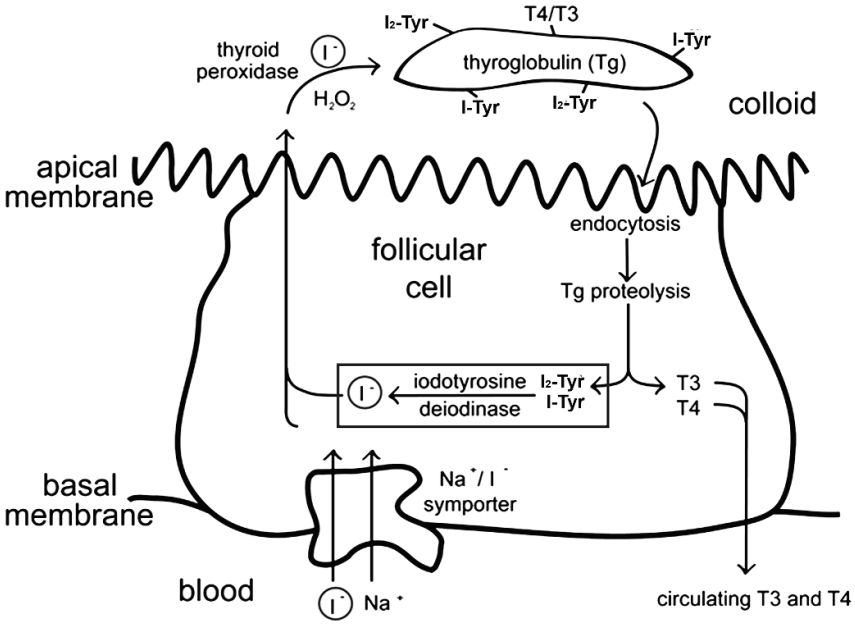

Figure 1-2. Iodide uptake and recycling by the thyroid follicular cell. Salvage of iodide from iodinated tyrosines is performed by iodotyrosine deiodinase. Figure is adapted from Rokita et al.[13]

1.3 Introduction to IYD

Deiodination of iodotyrosines was first demonstrated over 60 years ago by using thyroid tissue slices.[16] Since then, IYD has been shown as the enzyme responsible for catalyzing this reaction in the mammalian thyroid gland (Fig. 1-3).[17,18] IYD reduces the carbon-halogen bond by means of a flavin (FMN) cofactor and such reductive dehalogenation is unprecedented in mammals.[19] Reductive dehalogenation is rare in aerobic life[20] and even more unusual in mammals. Only one other enzyme, iodothyronine deiodinase (ID), also critical for thyroid function catalyzes this process in mammals. ID catalyzes deiodination of THs and regulates their effects.[21]

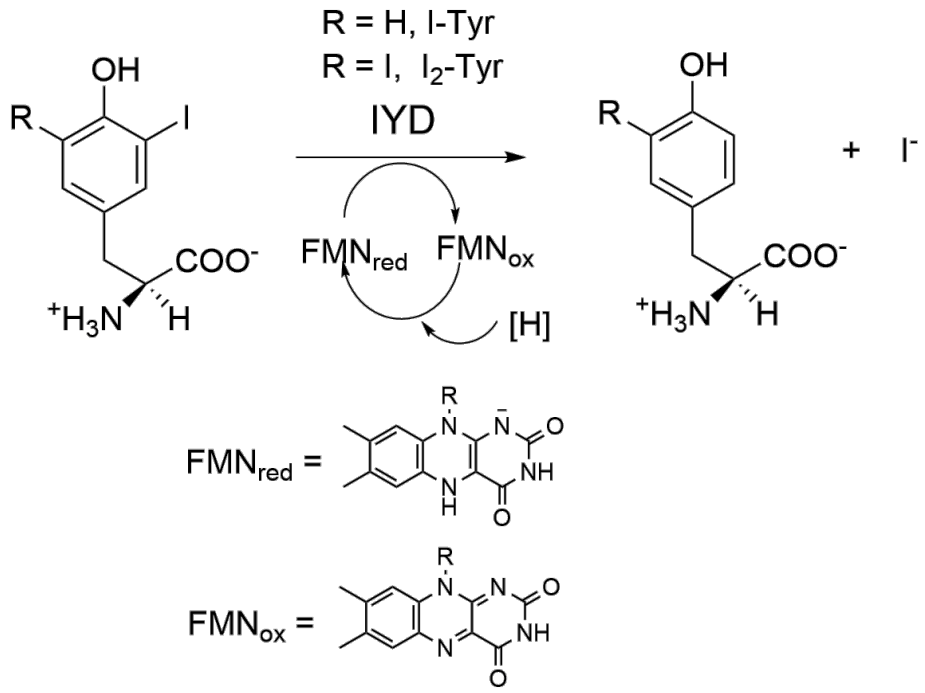

Figure 1-3. Deiodination of I-Tyr and I_2-Tyr catalyzed by IYD.

The origins of these two enzymes are quite different despite their ability to catalyze similar reactions. Unlike IYD that relies on a flavin cofactor, ID catalyzes deiodination of THs through an active site selenocysteine. ID belongs to the thioredoxin superfamily[22] while IYD on the other hand belongs to the nitro-FMN reductase superfamily (previously referred to as the NADH oxidase/flavin reductase superfamily).[23]

Sequence analysis of mammalian IYDs revealed the presence of an N-terminal membrane anchor, an intermediate domain that displays higher variability and lacks homology to known protein domains and a C-terminal catalytic domain that is well conserved and displays homology to members of the nitro-FMN reductase superfamily.[23] Although many nitroreductases are able to utilize reduced nicotinamides (NADH or NADPH) directly as co-substrates,[24] IYD is thought to derive its reducing equivalents in vivo from NADPH indirectly through an intermediate reductase.[17,25]

Figure 1-4. The N-terminal membrane domain (A), intermediate domain (B) and catalytic domain (C) of mouse IYD. Amino acid numbers for the predicted domain boundaries are indicated above. Figure is adapted from Friedman et al.[23]

1.4 Isolation and expression of IYD

The first successful attempt to isolate IYD from bovine thyroid microsomes was reported 47 years ago for which an enzymatic treatment (steapsin lipase) or detergent was used to solubilize the enzyme. This was followed by ammonium sulfate precipitation and multiple chromatographic separation procedures.[18] The protocol was tedious and hard to replicate.[23] Between 2004 and 2008, heterologous expression of IYD was reported in mammalian cell cultures.[14,23,25] A membrane domain truncated (ΔTM, amino acids 2-33), C-terminally His tagged version of mouse (mm) IYD (mmIYD ΔTM, His_6) was expressed in a soluble form in mammalian cell cultures.[25] It retained catalytic activity with dithionite as reductant but activity with NADPH was suppressed indicating the membrane anchor might be essential for interaction with the putative reductase.[25] Further truncation of mmIYD up to its catalytic domain resulted in insoluble and catalytically inactive enzyme.[26] Sufficient yield of soluble enzyme from mammalian cultures remained elusive. Finally, expression of mmIYD (ΔTM, His_6) in yeast (*Pichia pastoris*) and insect cell (Sf9) cultures resulted in soluble and catalytically active IYD in sufficient quantities to allow purification to homogeneity by Ni^{2+} affinity chromatography.[27] Sf9 expression provided higher yields of enzyme than yeast expression enabling crystallographic studies on the enzyme.[28]

Eukaryotic cell lines are more difficult to grow, maintain and perform routine genetic manipulations. Therefore, an alternative prokaryotic expression system for IYD was desired. Initial attempts at expressing mmIYD (ΔTM, His_6) in *E. coli* yielded insoluble protein despite attempting N-terminal fusions with multiple protein tags known to enhance solubility.[29,30] Additional mutations of Cys residues to Ala along with an N-terminal thioredoxin fusion finally provided soluble expression of a catalytically active IYD in *E. coli* in sufficient quantities to enable crystallographic studies as reported in 2012.[29] Subsequently, *E. coli* expression of N-terminal SUMO fusion of human (hs) IYD (ΔTM, His_6) as well as IYD homologs from other organisms such as zebrafish, honeybee and lancelet yielded soluble protein in sufficient yields to enable purification and characterization of the catalytic activity without requiring additional mutations at Cys residues.[30,31]

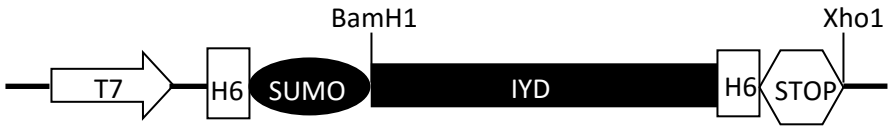

Figure 1-5. Design of SUMO fusion expression vector pSMT3[32] with the IYD gene. T7 is the promoter for expression and H6 is the His_6 affinity tag for Ni^{2+} affinity purification. BamH1 and Xho1 are restriction sites used to clone IYD genes into the vector. Figure is adapted from Buss J. M., PhD Dissertation.[30]

1.5 Relationship between IYD and other superfamily members

The nitro-FMN reductase superfamily is comprised of bacterial flavoproteins that catalyze diverse redox reactions of varied substrates including the reduction of nitroaromatics. IYD is the only known mammalian member of this superfamily.[24] The crystal structure of mmIYD reported in 2009 resulted in major advances in understanding

enzyme-substrate interactions and its relationship to other members within the nitro-FMN reductase superfamily.[28] Similar to other members of the superfamily, IYD exhibits an α-β fold with a dimer comprised of identical subunits that display domain swapping and an extensive dimer interface that includes the active site of the enzyme (Fig. 1-6).

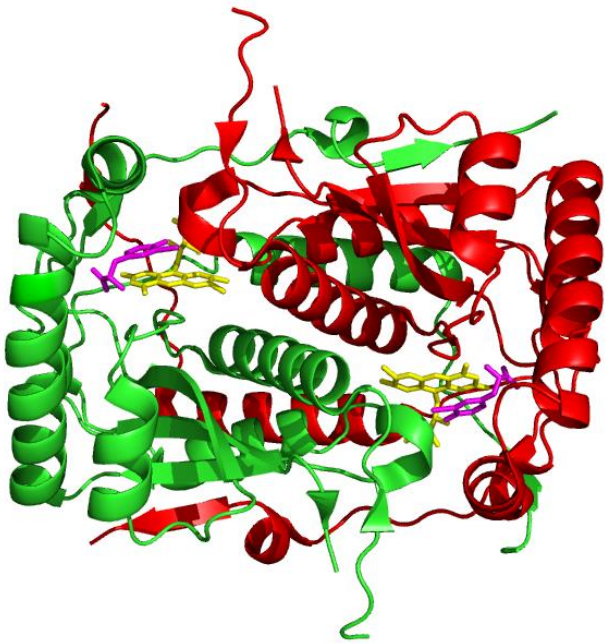

Figure 1-6. The co-crystal structure of mmIYD·I-Tyr (PBD ID 3GFD). The two identical polypeptides are indicated in green and red, respectively. FMN and I-Tyr in the active sites are indicated in yellow and purple respectively.

Within the superfamily, IYD is structurally most similar to BluB, a flavin destructase that degrades FMN to produce 5,6-dimethylbenzimidazole, the lower ligand of vitamin B_{12}.[33] Both IYD and BluB form a distinct subclass within the superfamily as their active site lids are derived from the same region of their primary structure.[28] The other two subclasses are represented by NADH oxidase (NOX) that accepts electrons from reduced nicotinamides for transfer to various acceptors[34] and by flavin reductase P (FRP) that provides reduced flavins for the luciferase activity.[35] The active site lid of these latter two enzymes is derived from different regions of their primary structure.[28] The presence of an

iodide salvage enzyme that functions in the mammalian thyroid within this superfamily of bacterial enzymes was not obvious and intriguing given its unusual recruitment of a flavin cofactor for dehalogenation.

1.6 IYD is more widespread than suggested by its known function of iodide salvage

IYD is primarily associated with iodide salvage and maintaining iodide homeostasis for thyroid hormone production.[13,36] However, the Basic Local Alignment and Search Tool (BLAST)[37] identified homologous proteins in diverse organisms including lower invertebrates such as arthropods (insects and crustaceans) and cnidarians (primitive marine organisms displaying radial symmetry including anemones, hydroids, jellies and corals).[30] Organisms belonging to these phyla lack the thyroid gland and are not known to synthesize THs endogenously. These findings suggested that the origins of IYD likely pre-date the evolution of TH signaling. Iodinated tyrosines have been proposed as possible evolutionary precursors of thyroid hormones[38] and IYD thus has the potential to serve as a marker to trace the origins of signaling via iodinated compounds.

A prerequisite for identifying additional IYDs throughout multiple phyla within different kingdoms in the tree of life is a robust set of determinants indicative of this enzymatic activity.

1.7 Specific Aims

Within the past decade, significant progress has been made in understanding the mechanism of dehalogenation promoted by IYD as well as interactions between the protein, FMN cofactor and substrate. Most studies published to date have been focused on mammalian enzymes. However, this enzyme likely has origins that pre-date its function in mammals. Identification of more primitive IYD homologs present the opportunity to

investigate its biological function prior to evolution of THs. Additionally, more primitive IYD homologs may exhibit broader substrate specificities enabling applications in bioremediation of halogenated aromatics. My first research objective was to investigate the origins of IYD within the nitro-FMN reductase superfamily through identification of signature residues that are markers for iodotyrosine deiodinase activity. My second research objective was to study the biological function of IYD in an invertebrate model organism not known to synthesize THs and hence having no known need for iodide salvage. This dissertation describes the following:

1) Identification of an amino acid signature that is a robust indicator of deiodinase activity derived from available crystallographic data, bioinformatics and phylogenetic analysis. Application of this signature sequence to test its predictive power through heterologous expression and catalytic characterization of IYD homologs across diverse phyla from eukaryotes to prokaryotes.

2) In vitro characterization of the IYD homolog from the model invertebrate *Drosophila melanogaster* not known to require iodide and lacking in TH synthesis. This includes probing the halotyrosine substrate specificity as well as site directed mutagenesis of highly conserved active site residues and its impact on substrate recognition and catalysis.

3) Investigation of the biological role of *Drosophila* IYD through loss of function studies including IYD gene silencing using RNA interference as well as gene editing using CRISPR/Cas9.

4) Probing the in vivo activity of *Drosophila* IYD though assessment of halotyrosine toxicity upon feeding halotyrosines to wild-type and IYD null mutants generated through gene editing by CRISPR/Cas9.

Chapter 2: Refining the signature sequence for IYD activity through heterologous expression and characterization of IYD homologs

2.1 Introduction

An enzyme superfamily generally represents a large group of proteins that are evolutionarily related to each other and have likely diverged from a common ancestor to perform specialized functions.[39] Members of a superfamily sharing structural similarities can catalyze diverse reactions and process diverse substrates.[40] Proteins of a superfamily can often be further organized into subgroups based on certain conserved sequence motifs and are thought to catalyze similar reactions. With ever increasing number of genomes from prokaryotes as well as eukaryotes being rapidly sequenced, proteins are being added to databases much faster than they can be analyzed for activity and therefore biochemical characterization of only a small fraction of all proteins from databases has been accomplished thus far. Annotation of newly added proteins is most often only predicted through computational methods. Automated annotation tools generally assign proteins to the correct superfamily. However, they are less accurate at assigning correct catalytic function.[41,42]

New protein sequences added to the nitro-FMN reductase superfamily are often broadly annotated as nitroreductases without accounting for the catalytic diversity of the members of this superfamily that includes iodotyrosine deiodinases (IYD), flavin destructases (BluB), flavin reductases (FRP) and NADH oxidases (NOX) among others as described in the previous chapter. Correct annotation of the plethora of new sequences will be crucial towards discovery of newer substrate specificities, potential applications in biomedical

research such as prodrug activation by nitroreductases in cancer therapy[24] and in bioremediation of halogenated aromatics.[19]

The co-crystal structure of mm IYD with its substrate I_2-Tyr identified key interactions between protein residues, substrate and cofactor and provided the opportunity to identify a signature sequence that should be diagnostic of IYD enzymes within the nitro-FMN reductase superfamily. The side chains of Glu-153, Tyr-157 and Lys-178 interact with substrate zwitterion[28] and are not present in BluB, the closest structural relative of IYD.[33] The amide nitrogen of Ala-126 donates a hydrogen bond to the substrate phenolate and is replaced by Gly-61 at an equivalent position in BluB. The OH group of Thr-235 hydrogen bonds with O^4 and $N^5$of FMN and a similar interaction is provided by Ser-167 in BluB. A side chain hydrogen bonding interaction with flavin N^5 has not yet been tested in other members of the superfamily.

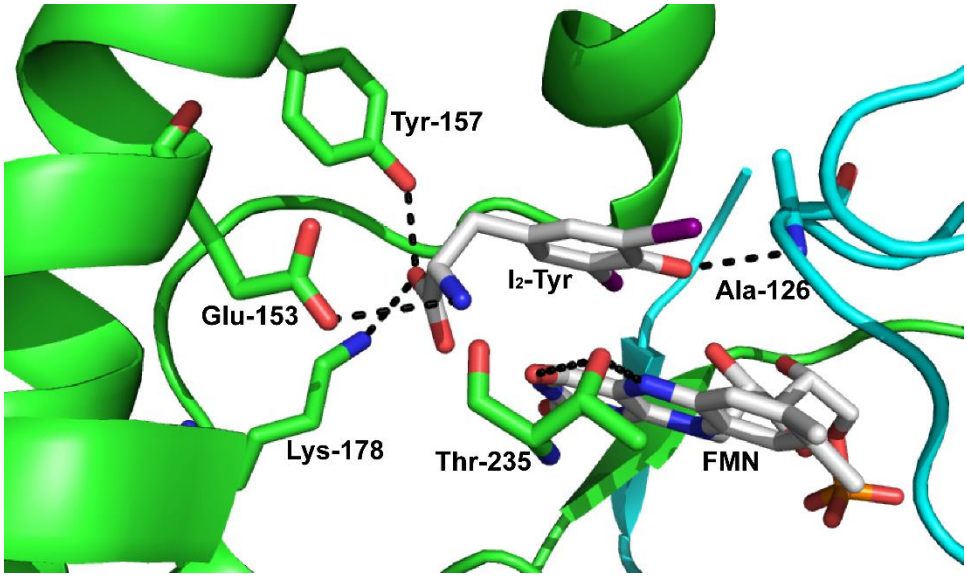

Figure 2-1. Active site of mmIYD indicating the flavin cofactor (FMN), diiodotyrosine (I_2-Tyr) and residues interacting with cofactor and substrate (PDB ID 3GH8).[28] The two identical polypeptides comprising native IYD assembly are indicated in green and cyan.

In this chapter we show that the residues equivalent to Glu-153, Tyr-157, Lys-178 and Thr-235 of mmIYD accurately identify deiodinase enzymes. The presence of IYD enzymes is more pervasive thoughout the tree of life than initially thought and includes diverse organisms across metazoa as well as prokaryotes not initially associated with the need for iodine conservation.

2.2 Experimental Procedures

2.2.1 Materials

E. coli expression optimized genes for homologous IYDs were synthesized by Blue Heron Biotechnology (Bothell, WA). The pSMT3 plasmid and the SUMO specific protease Ulp1 expression plasmid (pET28b-Ulp1)[32] were kindly provided by Dr. C. Lima (Memorial Sloan-Kettering Institute, New York, NY). The BluB enzyme was kindly provided by Dr. M. E. Taga (UC Berkeley, Berkeley, CA). The SsuE enzyme was kindly provided by Dr. H. R. Ellis (Auburn University, Auburn, Alabama). Expression plasmids (pET28) encoding IYD homologs from *Streptomyces sp. HPH0547* (stIYD), *Vibrio Harveyi CAIM 1792* (vhIYD) and *Thermotoga neapolitana* (tnIYD) were provided by Dr. Janine Copp (University of British Columbia, Vancouver). All DNA primers were synthesized by IDT (Coralville, IA). The pET24a plasmid and the Rosetta 2 DE3 *E.coli* cells were obtained from Novagen (Darmstadt, Germany). Phusion high fidelity DNA polymerase and its buffer, PageRuler broad range unstained protein standards, GeneRuler 1Kb plus DNA standards, Genejet plasmid miniprep kit, gel extraction kit, His-Pur Ni-NTA resin and ScintiSafe Plus 50% scintillation cocktail were obtained from Thermo-Fisher Scientific (Waltham, MA). All restriction endonuclease enzymes were obtained

from NEB (Ipswich, MA). T4 DNA ligase was obtained from either NEB or Promega (Madison, WI). Genehogs electro-competent *E. coli* cells were obtained from Invitrogen (Carlsbad, CA). The Sephacryl S-200 HR resin for size exclusion chromatography (SEC) was obtained from GE Healthcare (Piscataway, NJ).

2.2.2 General methods

Plasmids were purified from cultures of single colonies of *E. coli* cells using the Genejet plasmid miniprep kit. All DNA PCR reactions were performed using an Eppendorf Mastercycler gradient (Hauppauge, NY). All restriction digestions for PCR products as well as plasmids were performed at 37 °C for 1 to 2 hrs. PCR products and linearized plasmids were purified and visualized through agarose gel electrophoresis performed at 100 V. DNA was extracted from agarose gels using the Genejet gel extraction kit. All ligation reactions were performed with T4 DNA ligase. Transformation of *E. coli* with desired plasmids was performed using the Eppendorf Eporator at 2500 V.

Protein purification was performed on an AKTA Prime FPLC instrument (GE Healthcare) for both Ni^{2+} affinity chromatography (1 ml/min) and size exclusion chromatography (SEC, 0.3-0.4 ml/min). Purified proteins were concentrated using an Amicon Ultra 15 centrifugal filter with a 10,000 molecular weight cut off (Millipore). Proteins were analyzed by SDS-PAGE at 195 V using 12% resolving and 5% stacking acrylamide layers. The ε_{280} used for protein concentration, theroretical pI and molecular weights of expressed IYD homologs were determined from their amino acid sequences by ExPASy ProtParam.[43] All UV-visible absorption measurements were recorded with a Hewlett-Packard 8453 spectrophotometer. The FMN occupancy per enzyme active site was determined by measuring its absorbance at 450 nm (ε_{450} 12500 M^{-1} cm^{-1})[44] and the

concentration of protein was determined from its absorbance at 280 nm (ε_{280} for the proteins is provided in Appendix Table A1) after subtracting the contribution from FMN absorbance calculated using an A_{280}/A_{450} of 1.57 measured using free FMN.

2.2.3 Identification, alignment and phylogenetic analysis of IYD homologs

Within metazoa, IYD homologs were identified through BLAST[37] searches using the mmIYD sequence as query and specifying the species representing each phylum within the search parameters. Representative prokaryotic IYDs were selected through broader searches within entire domains (eubacteria and archaea). Sequence alignments of selected IYD homologs were performed using MUSCLE.[45] Membrane spanning regions were predicted using TMHMM.[46] Catalytic domains of IYDs (excluding N-terminal membrane regions and intermediate domains if any) were used for phylogenetic analysis.

Table 2-1. Protein accession numbers for the IYD protein homologs deposited in the NCBI database.

Organism	Protein accession
Daphnia pulex (dpIYD)	EFX90111.1
Nematostella vectensis (nvIYD)	XP_001633169.1
Hydra magnipapillata (hmIYD)	XP_002164528.1
Haliscomenobacter hydrossis DSM 1100 (hhIYD)	YP_004447048.1
Pyrococcus furiosus (pfIYD)	WP_011012802.1

2.2.4 Cloning of homologous IYDs

Expression optimized genes were synthesized with a C-terminal His_6 tag followed by a stop codon and flanking 5' Bam HI and 3' XhoI restriction sites used for subcloning into the pSMT3 vector to create N-terminal SUMO fusions. Forward primer 5'-AATT-AATCATATGCGTGTCCTCGAACTTGC-3' and reverse primer 5'-AATTAAT-CTCGAGTTAGTGATGATGATG-3' containing NdeI and XhoI restriction sites

respectively (underlined) were used to amplify the pfuIYD gene from the pfuIYD-pSMT3 vector template using Phusion DNA polymerase as per manufacturer recommended protocol. The restriction sites underlined above were used to subclone the pfuIYD gene into the pET24a vector for expression without N-terminal SUMO fusion.

2.2.5 Expression and purification of homologous IYDs

Rosetta 2(DE3) *E. coli* cells were transformed with the appropriate plasmid and single colonies were picked to inoculate a 20 ml starter culture of Luria LB media containing kanamycin (50 μg/ml) and chloramphenicol (34 μg/ml) followed by incubation at 37 °C with agitation for 13-15 hrs. The starter culture was diluted 50-fold in 1 L of Luria LB media and incubated with agitation at 37 °C until an OD_{600} of 0.6-0.8 was reached. Protein expression was induced by addition of IPTG under optimized conditions (Table 2-2).

Table 2-2. Induction conditions during expression of IYD homologs.

IYD homolog	IPTG (μM)	Temp (°C)	Time (hrs)
dpIYD	50	16	12
nvIYD	400	18	4
hmIYD	20	16	12
hhIYD	400	18	4
pfIYD[a]	20	16	12-14
stIYD	25	16	12-14
vhIYD[b]	25	16	12-14
tnIYD	25	16	12-14

Expression and purification performed by [a]Nattha Ingavat and [b]Jamie Alley.

Cells were then harvested by centrifugation, frozen in liquid nitrogen and stored at -80 °C. Prior to purification, cells were resuspended in lysis buffer (500 mM NaCl, 0.1 mM DTT, 10% glycerol and 50 mM sodium phosphate, pH 8.0) and lysed by 3 passages through a

French press cell at a sustained pressure of 1000 psi or by 3 passages through an Emulsiflex C3 homogenizer (Avestin) at a pulse pressure of 15,000 psi. The lysates were centrifuged at 40,000 × g for 2 hrs to separate the cell debris. Initial purification of SUMO-IYD fusions and non-SUMO tagged proteins was performed using Ni^{2+} affinity chromatography. The supernatant (40 to 60 ml) was loaded on HisPur Ni-NTA column (5 ml) pre equilibrated with lysis buffer and subsequently washed with 5 column volumes of lysis buffer containing 20 mM and 60 mM imidazole and 2 to 3 column volumes of lysis buffer containing 100 mM imidazole. The proteins were eluted with lysis buffer containing 250 mM imidazole. Purified SUMO-IYD fusions in elution buffer were treated with Ulp1 (approximately 1:200 w/w) for 12 hrs.

The bacterial hhIYD was separated from SUMO by using a second Ni^{2+} affinity column (Hi-trap, 1 ml, GE Healthcare). Protein was dialyzed with lysis buffer to remove imidazole prior second Ni^{2+} affinity purification. After loading, column was washed with 10 ml of lysis buffer containing 20 mM and 60 mM imidazole followed by 5 ml of lysis buffer containing 100 mM imidazole. The protein was eluted with lysis buffer containing 250 mM imidazole. Fractions were analyzed by SDS-PAGE and SUMO containing fractions were excluded.

Homologs dpIYD, nvIYD and hmIYD were concentrated to 1-1.5 ml and further purified by SEC (200-220 ml resin per column) using buffer containing with 300 mM NaCl, 1 mM DTT, 10% glycerol and 50 mM sodium phosphate pH 7.4. Similarly, homologs tnIYD and stIYD were further purified by SEC using buffer containing 100 mM NaCl, 0.5 mM TCEP, 15% glycerol and 50 mM sodium phosphate pH 7.4. A third purification step using a Hi-trap Ni^{2+} affinity column as described above was necessary for

dpIYD. The amino acid sequence of the active dpIYD fragment was confirmed by N-terminal protein sequencing (JHMI core facility). The pfuIYD homolog was expressed without the SUMO tag and purified by Nattha Ingavat who kindly provided the enzyme for activity analysis. The vhIYD homolog was expressed, purified and assayed for catalytic activity by Jamie Alley.

2.2.6 Analysis of hmIYD

The presence of hmIYD in protein isolates previously assayed for deiodinase activity was confirmed by western blotting using an anti-His_6 mouse monoclonal primary antibody (Novagen) and a goat anti-mouse secondary antibody conjugated to alkaline phosphatase (AP, Novagen). Development of the western blot was performed using the AP detection reagent kit (Novagen) as per the manufacturer's recommended protocol. The total protein concentration was determined by a BCA assay as per the manufacturer's protocol and the percentage of hmIYD was estimated by densitometric evaluation (ImageQuant TL, GE Healthcare) of an SDS-PAGE protein gel visualized by coomassie staining. The concentration of hmIYD was estimated from the observed average molecular weight of 33.5 kDa (See results).

2.2.7 Deiodinase activity assays

The deiodination rate of radiolabeled [^{125}I]-I_2-Tyr was determined by quantifying release of [^{125}I]-iodide as previously described.[23,36] Briefly, assay mixtures contained 200 mM KCl, 50 mM 2-mercaptoethanol, 0.05 mM methimazole, 0.033 mM FMN and 100 mM potassium phosphate pH 7.4 (unless specified otherwise). The total I_2-Tyr and enzyme concentrations was optimized for each homolog. Each reaction tube contained radiolabeled substrate [^{125}I]-I_2-Tyr (15,000 – 20,000 cpm). Enzyme reactions were initiated by addition

of 10% sodium dithionite dissolved in 5% aqueous sodium bicarbonate (100 µl) to a final volume of 1 ml and incubated at 25 °C unless specified otherwise. Reactions were quenched by addition of 0.1% I_2-Tyr in 0.1 N sodium hydroxide (100 µl). An aliquot from each reaction mixture (250 µl) was transferred to a scintillation vial containing 4.75 ml of 10% acetic acid (S). The remaining reaction mixture (850 µl) was loaded onto AG 50W-X8 cation exchange resin (Bio-Rad) pre-equilibrated with 10% acetic acid. The eluate was collected in scintillation vials and additional 10% acetic acid (4.15 ml) was added to resin and collected in same vial (A). The resin was further washed with 10% acetic acid (5 ml) and eluate was collected in a new scintillation vial (B). Scintillation cocktail (12-15 ml) was added to samples S, A and B and the radioactivity (cpm) in each sample was quantified using a Tri-Carb 2910TR scintillation counter (Perkin Elmer). The fraction of radioactivity released was determined from equation 2-1. The initial rate (V) is determined from equation 2-2 where F0 represents the background radioactivity from a control reaction performed without addition of enzyme. The K_m and V_{max} parameters for each IYD homolog were determined by fitting the initial rate (nmol hr^{-1}) against substrate concentration [S] (µM) to the Michaelis-Menten equation (equation 2-3) using Origin 6.0 or 7.0.

$$F = \frac{\frac{cpmA + cpmB}{0.85}}{\frac{cpmS}{0.25}} \qquad \text{Equation } 2-1$$

$$V = (F - F0) \times \left(2 \times \frac{60\ (\text{mins hr}^{-1})}{\text{Time (mins)}} \times [S]\right) \qquad \text{Equation } 2-2$$

$$V = \frac{V_{max}\ [S]}{K_m + [S]} \qquad \text{Equation } 2-3$$

2.2.8 BluB activity assay

BluB was tested for native activity using the FMN depletion assay as previously described in the presence of an FMN reductase enzyme, SsuE and NADPH.[47] The depletion of added FMN substrate was monitored at 445 nm. FMN depletion was confirmed in the presence of BluB (Appendix, Fig. A1)

2.3 Results

2.3.1 Selection and sequence analysis of IYD homologs

We identified sequences homologous to mmIYD within the nitro-FMN reductase superfamily from a wide range of organisms representing diverse phyla within the animal kingdom as well as multiple prokaryotic IYD homologs from eubacteria as well as archaea. Many putative IYDs within the animal kingdom were annotated as such, however a few homologs from invertebrates lacked annotation. Prokaryotic IYD homologs from bacteria were mostly annotated as nitroreductases or sometimes more generally as oxidoreductases. IYD homologs from thermophilic archaea and bacteria were annotated as NADH oxidases.

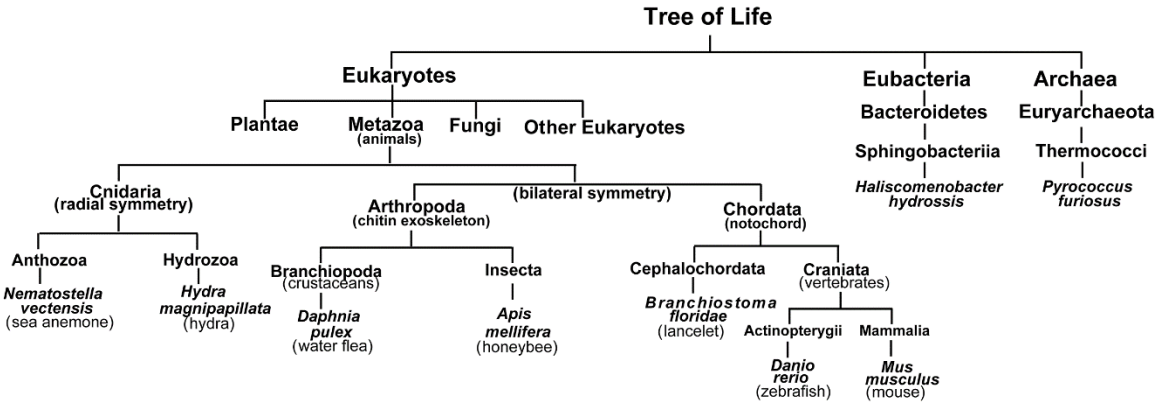

Figure 2-2. The tree of life indicating the diversity of organisms from which IYD homologs were chosen for expression. Phylum and class were obtained from the NCBI taxonomy browser.[48] Branches are not drawn to scale.

IYD homologs were selected from organisms representing distinct phyla within the animal kingdom as well as bacterial and archaeal representatives (Fig. 2-2). Organisms widely studied as genetic models were chosen whenever possible. IYD homologs from some lower eukaryotes and prokaryotes lacked predicted membrane anchors while the archaeal IYD homolog completely lacked the intermediate domain in addition to the membrane anchor.

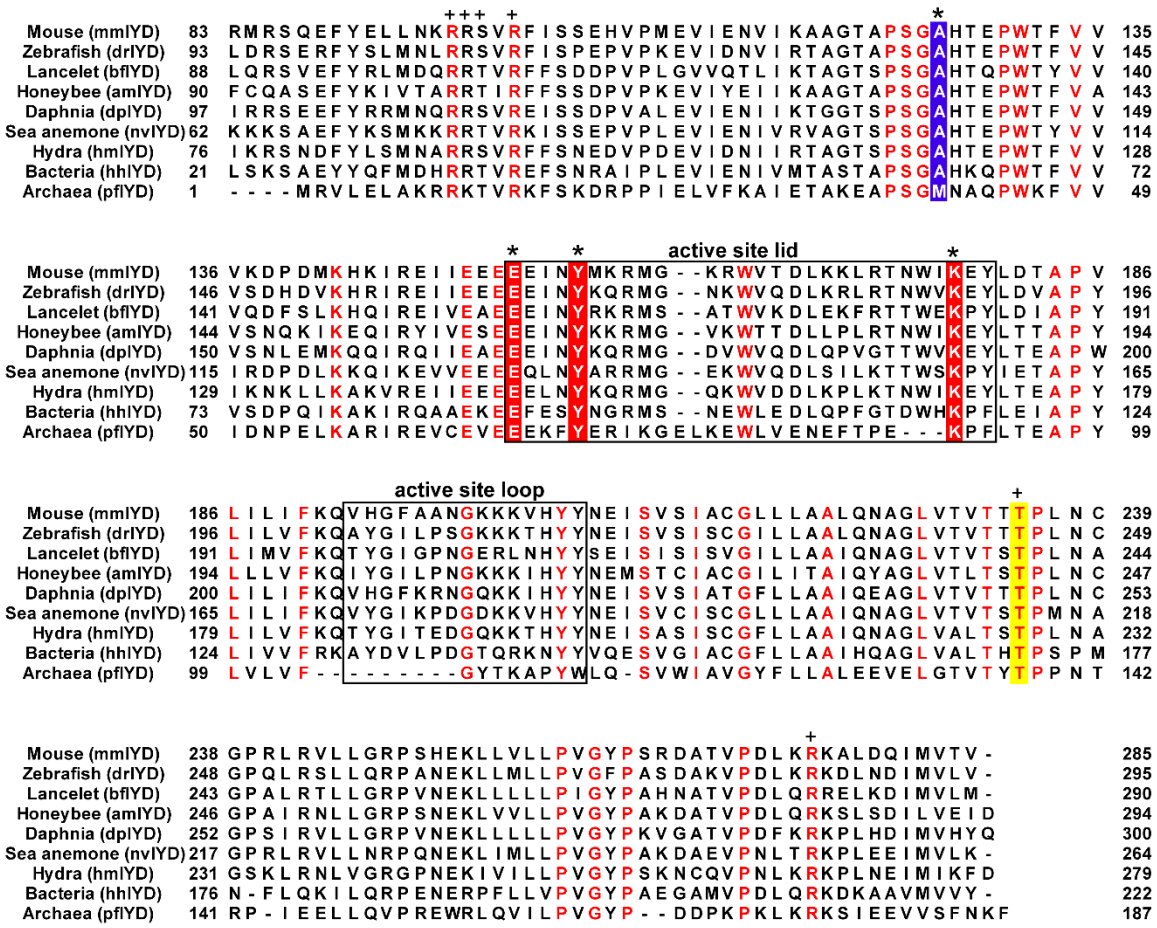

Figure 2-3. Multiple sequence alignment of catalytic domains of representative IYD homologs from selected phyla within the Metazoa kingdom and Prokaryota group. Numbering of amino acid residues is indicated on the left and right of the alignment. Residues in red (or white) are fully conserved. Substrate and FMN coordination residues are indicated with (*) and (+) respectively.[28] Columns highlighted in red indicate substrate zwitterion coordination residues. The column highlighted in blue indicates the residue that coordinates with substrate phenolate through a backbone amide and the column highlighted in yellow indicates the residue interacting with O^4 and N^5 positions of FMN.

All selected homologs showed conservation of the active site lid residues that are characteristic of the IYD subgroup within the superfamily and Thr interacting with FMN cofactor. All the homologs except pfIYD showed conservation of the Ala interacting with substrate phenolate group through amide backbone (Fig. 2-3).

2.3.2 Selection of diverse prokaryotic IYD homologs

The entire Nitro-FMN reductase superfamily of proteins was analyzed by Dr. Akiva (Dr. Babbit Lab, University of California at San Francisco) using Cytoscape[49] to generate sequence similarity networks (SSN).[50]

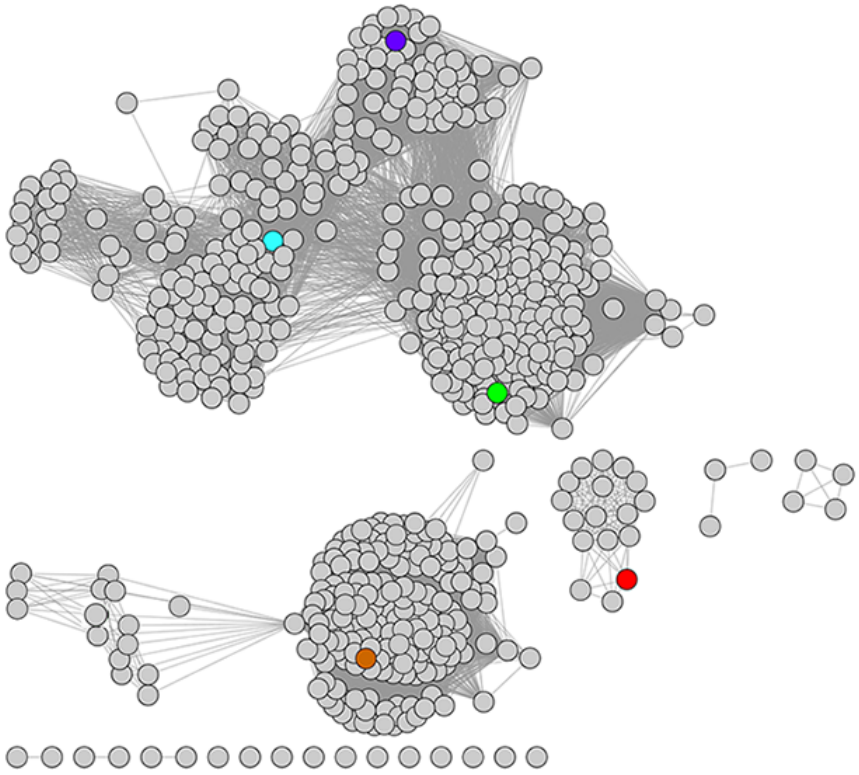

Figure 2-4. Cytoscape analysis of the IYD subgroup within the nitro-FMN reductase superfamily using a stringency threshold value of e^{-70}. Homologs mmIYD and hhIYD are indicated in brown and cyan respectively while stIYD, tnIYD and vhIYD are indicated in blue, red and green respectively. Analysis provided by Dr. Janine Copp (UBC).

Within the superfamily, the proteins belonging to the cluster of IYD enzymes were further analyzed by Dr. Janine Copp (University of British Columbia, Vancouver) (Fig. 2-4). Genes for prokaryotic IYD homologs from *Streptomyces sp. HPH0547* (stIYD, Uniprot S3BWA0), *Vibrio harveyi CAIM 1792* (vhIYD, Uniprot M7QQY2) and *Thermotoga neapolitana* (tnIYD, Uniprot B9K712) belonging to different sub-clusters within the IYD cluster were provided by Dr. Janine Copp for characterization of deiodinase activity. Active site residues and Thr interacting with FMN are also conserved in these prokaryotic IYD homologs (Fig. 2-5)

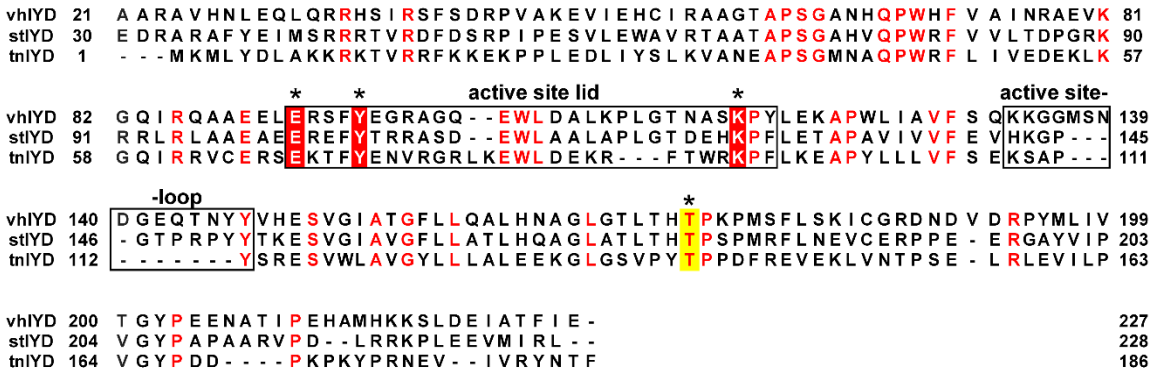

Figure 2-5. Multiple sequence alignment of catalytic domains from prokaryotic IYD homologs chosen for analysis of catalytic activity. Numbering of amino acid residues is indicated on the left and right of the alignment. Residues in red (or white) are fully conserved. Conserved residues thought to be diagnostic of IYD activity are indicated with (*). Columns highlighted in red indicate residues that coordinate the substrate zwitterion and the column highlighted in yellow indicates residue interacting with O^4 and N^5 positions of FMN.

2.3.3 Expression and purification of IYD homologs

All selected proteins were initially expressed as N-terminal SUMO fusions to enhance their solubility. Final yields of nvIYD and hhIYD ranged from 3 to 5 mg/L. Expression of dpIYD and hmIYD resulted in most of the protein being sequestered in inclusion bodies with yields lower than 1 mg/L. N-terminal sequencing of the protein isolated from dpIYD

purification revealed an N-terminus with sequence had been truncated in situ by 55 amino acids from the N-terminus (Appendix Fig. A2 and Fig. A3). Multiple proteins were captured from the soluble fraction of SUMO-hmIYD after Ni^{2+} affinity chromatography. SUMO-hmIYD was only partially processed by the SUMO specific protease Ulp1 (ca. 55%) indicating that the protein may be misfolded. Ni^{2+} affinity purification of SUMO-hmIYD was repeated and the proteins captured were further purified by SEC without Ulp1 digestion. The catalytically active fraction isolated by SEC consisted of multiple proteins some of which were derived from SUMO-hmIYD as determined by western blotting and coomassie staining (Appendix Fig. A4). A truncated SUMO-hmIYD (32 to 35 kD bands in western blot) comprising ca. 15% of all proteins detected in the catalytically active fraction was likely responsible for deiodination activity since full length SUMO-hmIYD (ca. 55 kDa, band in western blot) isolated separately in another SEC fraction did not display deiodinase activity. The sequence for hmIYD has since been revised in the NCBI database and differs from the expressed sequence in the first 42 amino acids which are part of the intermediate domain. This could explain the difficulties in expression of hmIYD.

The archaeal pfIYD homolog was initially expressed as a SUMO fusion protein. However, incomplete processing of the SUMO-pfIYD fusion following Ulp1 treatment indicated likely misfolding of the SUMO fusion protein. Subsequently, pfIYD was expressed without the SUMO tag and further purified by Nattha Ingavat to a final yield of about 7 mg/L. Other prokaryotic IYDs received from our collaborator Dr. Janine Copp were also expressed without SUMO tag. The protein yields of stIYD and tnIYD were high at 14 mg/L and 24 mg/L while vhIYD was purified by Jamie Alley at a very high yield of 59 mg/L.

All IYD homologs except hmIYD were purified to homogeneity (Figure 2-6). IYD homologs used for activity analysis contained FMN at occupancy greater than 87% other than nvIYD, hmIYD and vhIYD that had lower FMN occupancies of 55%, 59% and 63% respectively.

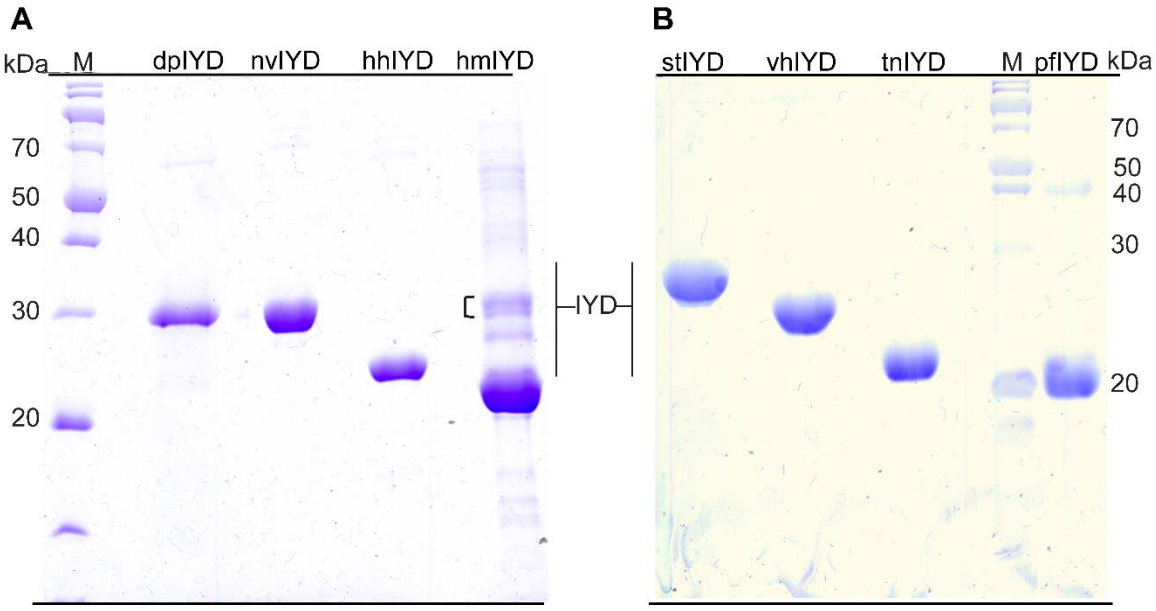

Figure 2-6. SDS-PAGE gel images of purified IYD homologs. IYD homologs from daphnia (dpIYD), sea anemone (nvIYD), *H. hydrossis* (hhIYD) and hydra (hmIYD) are shown in (A). Catalytically active fraction of hmIYD is indicated with brackets. IYD homologs from *Streptomyces Sp. HPH 0547, V. harveyi* (vhIYD), *T. neapolitana* and *P. furiosus* (pfIYD) are indicated in (B). M is marker lane with molecular weights of protein markers indicated on the left in kDa.

2.3.4 Deiodination activity of IYD homologs

All IYD homologs selected for catalytic activity analysis displayed the ability to turnover I_2-Tyr. Their k_{cat}/K_m values ranged between 0.12 min^{-1} μM^{-1} to 3.5 min^{-1} μM^{-1} (Table 2-3). The K_m values ranged between 1 µM and 105 µM while the k_{cat} values ranged 0.4 min^{-1} to 32 min^{-1}. The sea anemone homolog displayed the highest k_{cat} and K_m values while the highest catalytic efficiencies were displayed by hmIYD and the thermophilic pfIYD (at 60 °C). IYD homolog tnIYD displayed lowest k_{cat} of 0.45 min^{-1} that consequently

resulted in the lowest k_{cat}/K_m. Catalytic activity of pfIYD was assayed at pH 6.8 since the protein was prone to precipitation at pH 7.4. This insolubility at pH 7.4 was thought to be due to proximity to the theoretical pI of 7.8 as estimated by ExPASy ProtParam.[43]

The closest structural relative of IYD within the superfamily (BluB) with an alternative FMN destructase activity did not exhibit any deiodinase activity above the detection limit with $K_m \geq 133$ μM and $k_{cat} \leq 0.06$ min^{-1} and a $k_{cat}/K_m \leq 5 \times 10^{-4}$ min^{-1} μM^{-1}. No BluB-like activity has been previously reported when reduced mmIYD is exposed to BluB's co-substrate, molecular oxygen.[19] In addition to lack of BluB-like activity, there is no precedence for IYD exhibiting catalytic activity similar to other members of the superfamily represented by FRP and NOX that can directly accept electrons from reduced nicotinamides.[25]

Table 2-3. Catalytic deiodination by IYD homologs.

Homolog	Temp (° C)	K_m (μM)	k_{cat} $(min)^{-1}$	k_{cat}/K_m $(min^{-1}$ $\mu M^{-1})$
Human (hsIYD)[a]	25	31 ± 6	12.5 ± 1	0.40 ± 0.05
Mouse (mmIYD)[b]	25	19 ± 3	6.9 ± 1.3	0.36 ± 0.09
Zebrafish (drIYD)[c]	25	8 ± 1	4.1 ± 0.2	0.51 ± 0.07
Lancelet (bfIYD)[c]	25	6 ± 3	7 ± 1	1.2 ± 0.6
Honeybee (amIYD)[c]	25	29 ± 7	8 ± 1	0.28 ± 0.07
Daphnia (dpIYD)[d]	25	7 ± 1	17.5 ± 0.8	2.5 ± 0.4
Sea anemone (nvIYD)[d]	25	105 ± 26	32 ± 4	0.30 ± 0.08
Hydra (hmIYD)[d]	25	4 ± 1	14 ± 1	3.5 ± 0.9
H. hydrossis (hhIYD)[d]	25	6.6 ± 0.9	5.4 ± 0.3	0.8 ± 0.1
Streptomyces Sp. (stIYD)[d]	25	7 ± 2	3.2 ± 0.3	0.5 ± 0.1
V. harveyi (vhIYD)[d,e]	25	11 ± 3	15 ± 1	1.4 ± 0.4
P. furiosus pfIYD[d,f]	25	1.0 ± 0.1	0.442 ± 0.008	0.4 ± 0.1
P. furiosus pfIYD[d,f]	60	1.8 ± 0.3	5.5 ± 0.2	3.1 ± 0.5
T. neapolitana (tnIYD)[d]	25	-	< 0.4	-
T. neapolitana (tnIYD)[d]	60	3.8 ± 0.7	0.45 ± 0.02	0.12 ± 0.02

[a-c] Parameters were previously reported.[29-31] [d]Determined from data illustrated in Appendix Fig. A5. Errors derive from least square fitting. [e]Assay performed by Jamie Alley. [f]Enzyme assays conducted at pH 6.8.

2.4 Discussion

All IYD homologs from eukaryotes exhibited deiodination activity with catalytic efficiencies distributed within an order of magnitude. Amongst the prokaryotic IYD homologs evaluated, all displayed deiodination rates within an order of magnitude with the exception of tnIYD that displays about 25 to 30-fold lower catalytic efficiency than the most efficient deiodinases. However, the catalytic efficiency of tnIYD is still within an order of magnitude of the average catalytic efficiency displayed by other enzymes. Increased k_{cat} at 60 °C compared to 25 °C is expected for IYD homologs from thermophilic (*T. neapolitana*) or hyperthermophilic (*P. furiosus*) organisms.

The key residues selected from the crystal structure of mmIYD (Glu-153, Tyr-157, Lys-178 and Thr-235) are indeed excellent predictors of deiodinase activity despite the large variations in sequence length (including gaps in the active site loop regions of prokaryotic thermophilic IYDs) and identity (Figs. 2-2 and 2-4) . We had initially included the Ala-126 (blue highlight Fig. 2-2) providing the backbone hydrogen bonding interaction as a key identifier of IYD. However, expanding our BLAST search with lower stringencies (expect values $>10^{-40}$) identified homologous proteins from thermophilic prokaryotes wherein the Ala-126 was replaced by a Met at the equivalent position. IYD homologs pfIYD and tnIYD, both having this Met residue instead of an Ala display deiodinase activity. Since this residue provides a backbone amide hydrogen bonding interaction to the phenolate of substrate, lack of conservation is not surprising since this interaction could still be maintained by another amino acid.

The phylogenetic analysis of amino acid sequences of IYD homologs, nitro-FMN reductases (FRP and NOX) and BluB (Fig. 2-5) indicates the divergent evolution of

representatives within this superfamily. Clearly, the thermophilic prokaryotic IYDs are most distant from IYDs of higher organisms and are more likely to exhibit some promiscuity. This may explain the relatively low k_{cat} and k_{cat}/K_m exhibited by tnIYD. Despite low sequence identity (19%), BluB and IYD share significant similarities with their active site lids derived from the same region of their primary structure.[28,33] The FMN cofactor in both BluB and IYD is primed for sequential one electron chemistry.[13,31] Inspite of these similarities, BluB does not catalyze deiodination of I_2-Tyr above our limits of detection.

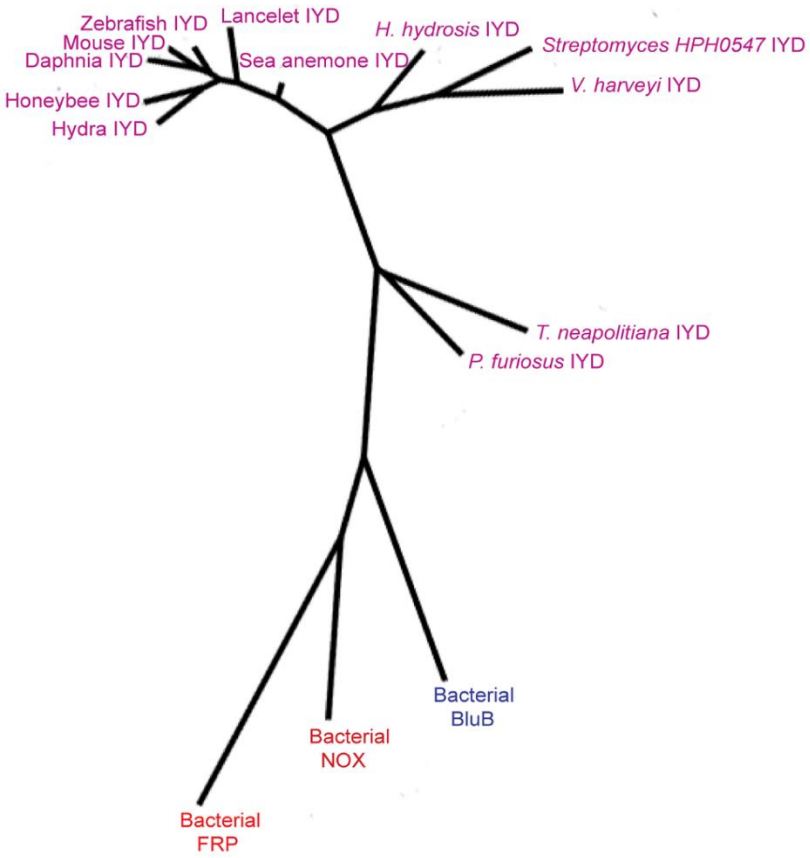

Figure 2-7. Phylogenetic analysis of the IYD subgroup proteins and other representatives of the nitro-FMN reductase superfamily. This phylogenetic tree was generated by phylogeny.fr through integration of MUSCLE,[45] PhyML,[51] and TreeDyn[52] for multiple sequence alignment, phylogenetic analysis and tree rendering respectively. Catalytic domains of IYD homologs and aligned regions of BluB (PDB ID 2ISJ), NOX (PDB ID 1NOX) and FRP (PDB ID 2BKJ) were used as input sequences. The branch lengths are proportional to the number of amino acid substitutions per aligned residue.

The presence of IYD was expected within chordata (Fig. 2-2) as organisms within this group have a thyroid gland or an endostyle and likely produce THs requiring salvage of iodide. The presence of this enzyme within arthropods represented by insects and crustaceans is surprising and very intriguing since there is no confirmed evidence for endogenous TH production[53] although some reports suggest that these may be obtained from external sources[53,54] and affect insect metabolism.[55-57] The requirement for iodine recycling and role of iodotyrosines in insects still remains elusive. Within Cnidaria, the role of IYD in hydra and sea anemone is unknown and no reports were found regarding effects of iodide, iodotyrosines or THs in these organisms. Amongst other cnidarians, the above mentioned compounds were shown to be capable of initiating strobilation (metamorphosis from polyp to medusa) of the jellyfish *Aurelia aurita*.[58] I_2-Tyr was shown to be the most potent stimulator of this change indicating that iodotyrosines could be evolutionary precursors of the THs as proposed previously.[38] In this case, IYD could be functioning in a regulatory capacity.

A regulatory or iodide salvage role cannot be ascribed to IYD homologs from prokaryotes although some bacterial strains have been shown to accumulate iodine.[59,60] The enzyme is likely involved in catabolism or detoxification enabling bacterial and archaeal species to adapt to environments containing iodinated compounds. In addition, chlorinated and brominated compounds might also be substrates for these prokaryotic IYDs since at least mmIYD has already been shown to dehalogenate bromo- and chlorotyrosine.[19]

No IYD homolog has been identified within fungi. BLAST search for IYD homologs in fungi yielded proteins with weak homology not showing conservation of any of the

signature residues previously described. Even the fungi *Neurospora crassa* that encodes several metabolic flavoproteins[61] did not contain an IYD gene. Plant proteins showed very poor homology to members from the nitro-FMN reductase superfamily.

2.5 Summary

The residues Glu-153, Tyr-157, Lys-178 coordinating the substrate zwitterion and residue Thr-235 coordinating the O^4 and N^5 of FMN cofactor as identified from the crystal structure of mmIYD are excellent predictors of deiodinase activity. IYD homologs from a diverse range of organisms ranging from prokaryotes (bacteria and archaea) to representatives of several phyla within metazoa displayed deiodinase activity in vitro with catalytic efficiencies ranging from 3.5 min^{-1} uM^{-1} to 0.12 min^{-1} uM^{-1}. Sequence identity between the most distant IYD homologs (mmIYD and tnIYD) was just 37% however, all the aforementioned residues were conserved. Our results should facilitate functional assignments for putative IYDs within the nitro-FMN reductase superfamily. The presence of IYD is more widespread than initially thought based on its known function of iodide salvage in the thyroid gland of mammals. Catalytically competent IYDs have also been identified in lower invertebrates and prokaryotes not known to require iodine for TH synthesis.

Chapter 3: Functional analysis of the *Drosophila* halotyrosine dehalogenase

3.1 Introduction

IYD is primarily associated with iodine salvage from I-Tyr and I_2-Tyr, the byproducts of TH synthesis.[13] Surprisingly, we identified functional IYDs even in non-chordate invertebrates such as arthropods (insects and crustaceans) and cnidarians (radially symmetrical lower invertebrates such as anemones and jelly fish) that lack a thyroid gland or endostyle and no conclusive evidence for endogenous TH synthesis by these organisms has been reported (Chapter 2).

Previously, IYD has also been shown to promote dehalogenation of Br-Tyr and Cl-Tyr in addition to I-Tyr in single turnover experiments.[19,62] As an iodine requirement for thyroid hormone synthesis is unlikely in non-chordates, the primary substrates for their IYDs could be other halotyrosines like Br-Tyr and Cl-Tyr. Halogenated tyrosines have been reported in scleroproteins of insects[63] and their formation would not be surprising since Tyr plays a primary role in formation of cross-linked proteins.[64] IYD within insects could potentially serve as a general halotyrosine dehalogenase enzyme or may exhibit a particular halotyrosine preference characteristic of its physiological substrate.

Previously, the IYD homolog from *Apis mellifera* (honeybee) was selected as the insect representative for expression and deiodinase activity testing[30] rather than from *Drosophila melanogaster*, an insect model organism used extensively to probe gene function, since the *Drosophila* IYD homolog showed the presence of a large N-terminal domain of unknown function in addition to the deiodinase domain. However, the deiodinase domain shows high conservation of all residues now shown to be robust predictors of IYD activity, including

Glu, Tyr and Lys interacting with the substrate and Thr interacting with the FMN cofactor (Fig. 3-1).

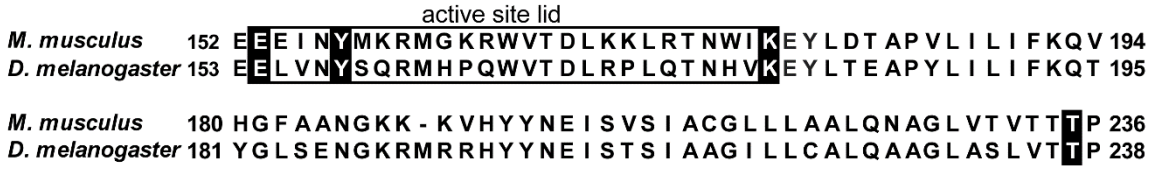

Figure 3-1. Multiple sequence alignment of catalytic regions of mouse and *Drosophila* IYD. Sequence alignment was performed using MUSCLE.[45] Numbering of amino acid residues is indicated on the left and right of the alignment. Conserved residues diagnostic of IYD activity are indicated in white.

In addition, alternative splicing of the IYD encoding gene annotated as CG6279 in the FlyBase is predicted to generate isoform B that lacks the additional N-terminal domain entirely (Appendix Fig. B1).[65] Isoform B of *Drosophila* IYD lacking membrane anchor (dmIYD) is therefore an ideal choice for detailed in vitro studies of an IYD homolog from a model non-chordate invertebrate species wherein the biological role of this enzyme is not yet known.

In this chapter, we examine the preferred halotyrosine substrate for dmIYD through the measurement of equilibrium binding and steady-state kinetics of turnover of I-Tyr, Br-Tyr Cl-Tyr and F-Tyr. In order to determine the relative importance of evolutionarily conserved active site Glu, Tyr and Lys residues coordinating the zwitterion of substrate, we expressed active site mutants of dmIYD and evaluated their impact on binding affinity and catalysis.

3.2 Experimental Procedures

3.2.1 Materials

Pfu Turbo DNA polymerase and its buffer were obtained from Agilent (Santa Clara, CA). 3-Bromo-L-tyrosine (95% pure) was purchased from AEchem Scientific Corporation (Naperville, IL). 3-Fluoro-L-tyrosine (97% pure) was purchased from Astatech Inc (Bristol, PA). 3-Iodo-L-tyrosine (97% pure) and 3, 5-diiodo-L-tyrosine (97% pure) were purchased from Acros Organics and Sigma respectively. Other materials used are as described in chapter 2 (2.2.1).

3.2.2 Cloning of dmIYD

IYD homolog from *Drosophila melanogaster* (isoform B, Accession NP_001163414.1, Flybase CG6279-PB) was truncated to remove an N-terminal transmembrane domain (amino acids 2-40) predicted by TMHMM.[46] This truncated gene sequence is designated as dmIYD. Codon optimized gene sequence for *E. coli* expression including a His_6 tag and stop codon at C-terminus was synthesized by Blue Heron Biotechnology (Appendix Fig. B2) and subcloned to generate N-terminal SUMO fusions in the pSMT3[32] plasmid. Forward primer 5'-AAACAAACATATGAAAACTTATAATTTAGATGAAC-3' and reverse primer 5'-AATTAATCTCGAGTTAGTGGTGATGGT-3'containing NdeI and XhoI restriction sites respectively (underlined) were used to amplify and subclone the dmIYD gene into the pET24a vector for subsequent expression without the SUMO tag.

3.2.3 Site-directed mutagenesis of dmIYD

The pET24a-dmIYD vector was employed as the template for site-directed mutagenesis using the Pfu turbo DNA polymerase. Primer 5'-GTAGAACAAGAACAGTTAG-TCAATTAC-3' and its reverse complement were used to generate the E154Q mutation.

The Y158F mutation was introduced by the primer 5'-AAATTGTAGAACAAGAAG-AATTAGTCAATTTCTCCCAACGTATGCAT-3' and its reverse complement while the K179Q mutation was introduced using primers 5'-ACCAATCATGTACAGGA-ATACTTAACC-3' and its reverse complement. The mutated codon is indicated in red. Thermocycling conditions for site-directed mutagenesis were performed as recommended by the Quikchange protocol (Agilent). Proper amplification of the mutation containing plasmids was confirmed by 1% agarose gel electrophoresis prior to their transformation into electro-competent Genehogs *E. coli* cells (Invitrogen). Plasmids were purified from cultures of single colonies and introduction of the mutations was confirmed by DNA sequencing (Genewiz).

3.2.4 Expression and purification of dmIYD enzymes

For protein expression, the appropriate plasmid was transformed into Rosetta 2 (DE3) *E. coli* cells (Novagen). All other expression conditions were similar to those described in chapter 2 (2.2.5) with the exception that 25 mM imidazole was added to the lysis buffer. Purification of dmIYD expressed without SUMO tag was performed as described for stIYD and tnIYD proteins in chapter 2 (2.2.5).

3.2.5 Binding affinity determination

The binding affinity of dmIYD and its derivatives was determined by fluorescence quenching of the protein-bound FMN cofactor (λ_{ex} of 450 nm and λ_{em} of 520 nm) upon addition of ligands as previously described.[19,31] Fluorescence intensity was measured using a Fluoromax-4 fluorescence spectrophotometer (Horiba Scientific). Protein solutions were prepared in buffer containing 200 mM potassium chloride and 100 mM potassium phosphate pH 7.4 to a final concentration of 3 μM enzyme as measured by associated FMN

(ε_{450} 12,500 M^{-1} cm^{-1}).[44] Observed fluorescence signal (F) was normalized through division by the initial fluorescence signal (no ligand added, F_0) and plotted against added ligand concentration (S). Equilibrium dissociation constant (K_d) values were determined by non-linear least squares fitting of the data to equation 3[47,66] using Origin 6.0. F is the observed fluorescence signal, F0 is the initial fluorescence signal in absence of ligand, ΔF represents total change in fluorescence intensity and E_t represents enzyme concentration.

$$\frac{F}{F0} = 1 + \frac{\Delta F}{F0}\left(\frac{(K_d + E_t + S) - \sqrt{(K_d + E_t + S)^2 - 4\,E_t S}}{2E_t}\right) \qquad \text{Equation 3}$$

3.2.6 Catalytic activity assay with ^{125}I-[I_2-Tyr]

The deiodination of radiolabeled ^{125}I-[I_2-Tyr] was performed as described in chapter 2 (2.2.7) to measure catalytic activity of dmIYD and its mutants.

3.2.7 Kinetics of monohalotyrosine turnover by dmIYD

Catalytic dehalogenation of 3-halotyrosines was measured by quantifying the Tyr produced in the presence of enzyme by reverse phase HPLC using an Agilent 1100 series instrumental setup. The reaction buffer comprised of 200 mM KCl and 100 mM potassium phosphate pH 7.4. The internal standard (IS) included was m-cresol (30 µM). Reaction was initiated by addition of 5% sodium dithionite in 5% sodium bicarbonate (100 µl) to a final reaction volume of 1 ml with incubation at 25 °C. Reaction was quenched by addition of 88% formic acid (50 µl) and the entire mixture was injected onto a Microsorb MV 300-5 C18 analytical column (Varian) connected to a manually packed C18 silica guard column. The solvent system comprised of 0.44% aqueous formic acid (A) and 0.44% formic acid in acetonitrile (B). Analytes were separated using a linear gradient from 0% B to 5% B over 10 mins and 5% B to 60% B over the next 15 mins. A brief linear drop from 60% B to 30% B was added over the next 2 mins to prevent drifting of the baseline that could

affect the internal standard peak. The column was washed with 95% B and re-equilibrated with 100% A prior to the next injection. All analytes were detected at 280 nm. The concentration of Tyr produced by dmIYD was be determined from a standard curve generated by spiking known concentrations of Tyr into reaction mixtures lacking the halotyrosine substrate. The ratio of the areas of Tyr signal to IS signal was plotted against the concentration of Tyr spiked into the reaction and the equation of the line was determined by linear regression (Appendix Fig. B3). This equation was used to determine the concentration of Tyr produced during enzymatic reactions. The kinetic parameters were determined by fitting the data to the Michaelis-Menten equation as previously described in chapter 2 (2.2.7). Br-Tyr contained Tyr as an impurity and the contribution of the Tyr impurity was subtracted from the product signal prior to data analysis based on the standard curve (Appendix Fig. B4). The Tyr impurity in commercially purchased Br-Tyr was determined to be 5 % molar equivalents.

3.3 Results

3.3.1 Expression and purification of dmIYD proteins

Similar to other IYD homologs, dmIYD was initially expressed as a SUMO fusion followed by cleavage of SUMO by its selective protease Ulp1 and separation of SUMO by SEC. This process provided relatively low protein yields of 4 to 7 mg/L. Subsequently, dmIYD was expressed in *E. coli* without the SUMO fusion to streamline purification and generate high yields of 8-11 mg/L. All dmIYD mutants were also expressed without a SUMO fusion in yields between 5-12 mg/L (Fig. 3-2). FMN occupancy in dmIYD and its mutants was relatively consistent between 82 to 90%.

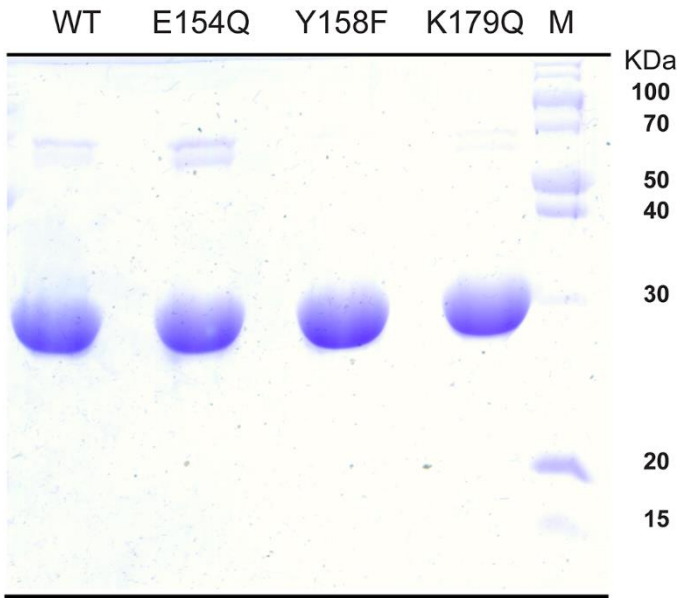

Figure 3-2. SDS-PAGE gel image of purified dmIYD and mutant proteins (~7.5 µg per well). M is marker lane with molecular weights of protein markers indicated on the right in kDa.

3.3.2 Substrate specificity of dmIYD.

The biological function of IYD in non-chordates has not been established and therefore the natural substrate for dmIYD could be a halotyrosine other than I-Tyr. To determine the specificity of dmIYD, the binding affinity and the kinetic parameters of steady-state turnover for each halotyrosine was measured. The enzyme exhibited little discrimination between I-Tyr, Br-Tyr and Cl-Tyr with K_d values lower than 1 µM (Fig. 3-3, Table 3-1). However, binding affinity of F-Tyr was significantly lower (25-fold) than that for other halotyrosines.

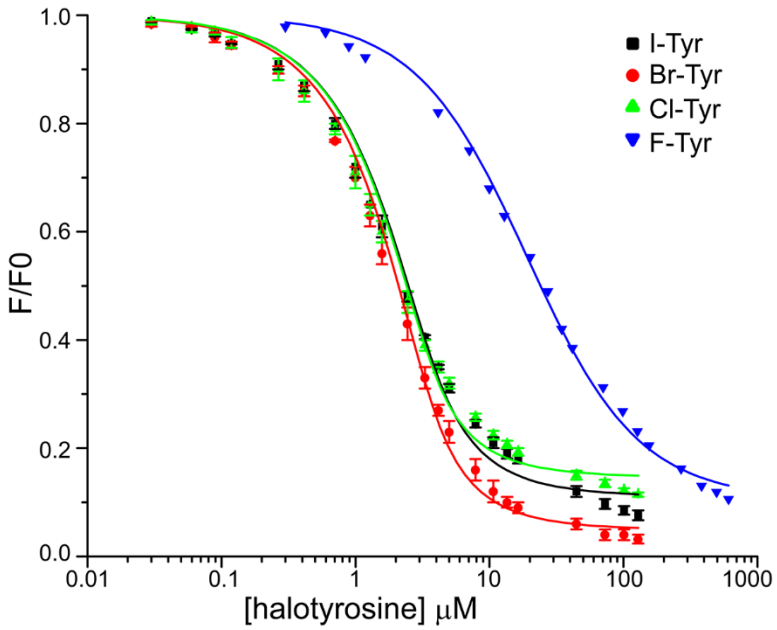

Figure 3-3. Quenching of FMN fluorescence is indicative of ligand binding in the active site of dmIYD. Each data point for I-Tyr, Br-Tyr and Cl-Tyr binding represents an average of 3 individual observations and error bars represents its standard deviation. F-Tyr binding curve is represented by a single data set. The K_d values are derived from the best fit of the data (solid lines) to equation 3.

The dehalogenation product (Tyr) from turnover by dmIYD was detected and quantified by HPLC (Figure 3-4, Appendix Fig. B5). The K_m was lowest for Br-Tyr and highest for Cl-Tyr while k_{cat} was highest for I-Tyr and lowest for Cl-Tyr (Table 3-1). Overall, catalytic efficiency of I-Tyr and Br-Tyr turnover was similar while turnover of Cl-Tyr was about 3.5-fold less efficient. We did not detect any Tyr formation above the limit of detection when F-Tyr was used as a substrate for dmIYD even in presence of 25-fold higher enzyme concentration and 6-fold longer incubation time (Table 3-1, Appendix Fig. B7).

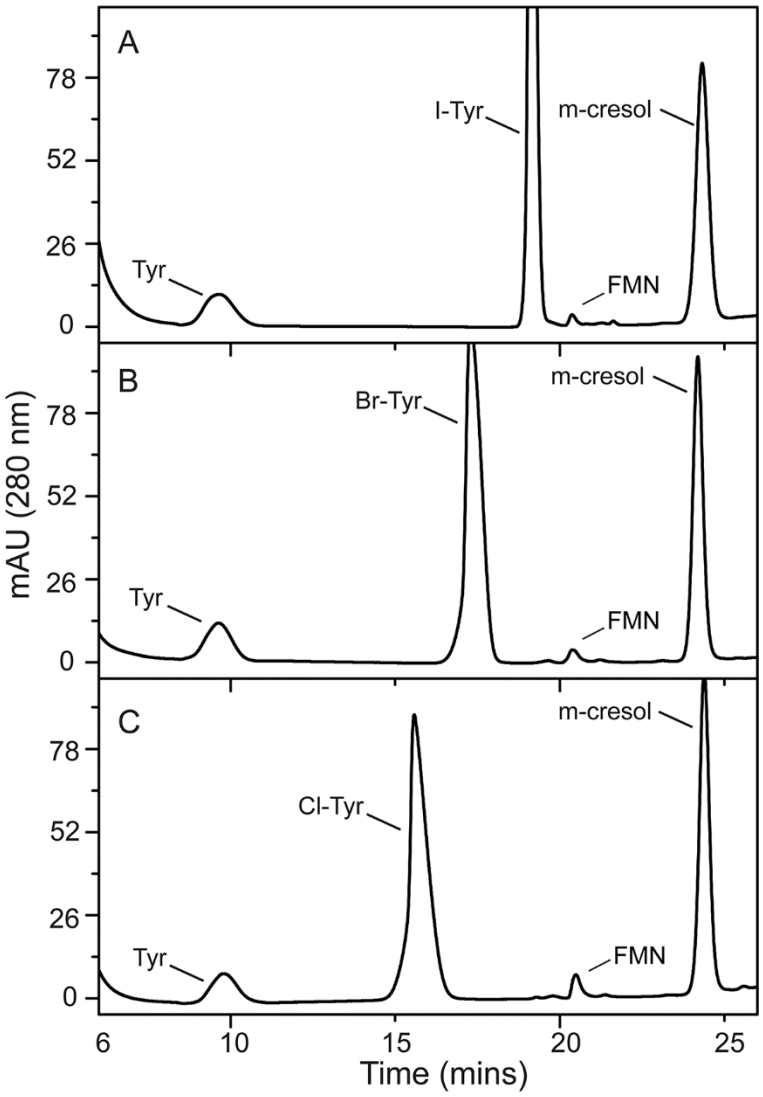

Figure 3-4. Detection of tyrosine after halotyrosine turnover by dmIYD using reverse phase HPLC. (A) I-Tyr (40 μM) was incubated with dmIYD (0.08 μM) for 20 mins. (B) Br-Tyr (40 μM) incubated with dmIYD (0.12 μM) for 20 mins. (C) Cl-Tyr (40 μM) incubated with dmIYD (0.2 μM) for 20 mins. Internal standard was m-cresol. For complete conditions refer to the experimental section. The FMN peak is derived from dmIYD.

Table 3-1. Binding and turnover of 3-halotyrosines by dmIYD

Substrate	K_d[a] (μM)	K_m[b] (μM)	k_{cat} (min^{-1})[b]	k_{cat}/K_m (min^{-1} μM^{-1})[b]
I-Tyr	0.62 ± 0.08	14 ± 4	7.3 ± 0.6	0.5 ± 0.2
Br-Tyr	0.47 ± 0.06	8 ± 2	4.2 ± 0.3	0.5 ± 0.1
Cl-Tyr	0.46 ± 0.08	21 ± 6	3.0 ± 0.3	0.14 ± 0.04
F-Tyr	17.6 ± 0.8	-	<0.002	-

[a]Determined from data illustrated Figure 3-2. [b]Determined from data illustrated in Appendix Fig. B6. Errors derive from least square fitting.

3.3.3 Contribution of zwitterion coordination residues towards substrate affinity

Three active site residues Glu-153, Tyr-157 and Lys-178 as seen in crystal structure of mmIYD·I_2-Tyr (Chapter 2, Fig. 2-1) coordinate the substrate zwitterion and include 3 of the 4 residues shown to be diagnostic of IYD activity in the previous chapter. These residues are likely unique to IYDs within the superfamily and critical for substrate recognition. Limited information regarding their relative importance was obtained previously from mutagenesis of the mouse IYD.[29] Interest in these residues was renewed when IYD was found to be pervasive throughout diverse organisms. Identifying their relative contribution will help to define the origins of this enzyme.

The dmIYD enzyme exhibits strong binding affinity towards I_2-Tyr with a K_d value of less than 1 μM. (Fig. 3-5 A, Table 3-2). Removal of the phenolic OH group coordinating the zwitterion by a Phe substitution (Y158F) results in a 7-fold decrease in affinity for I_2-Tyr. Replacement of the carboxylate group coordinating the zwitterion with an amide group (E154Q) results in a 30-fold decrease in affinity for I_2-Tyr while replacing the charged ammonium with an amide (K179Q) is most deleterious towards binding affinity and results in a 45-fold decrease in affinity for I_2-Tyr. While the K_d displayed by dmIYD

for I_2-Tyr and I-Tyr is very similar, each of the mutant enzymes display a 7- to 10-fold lower affinity for I-Tyr compared to I_2-Tyr (Fig. 3-5 B, Table 3-2).

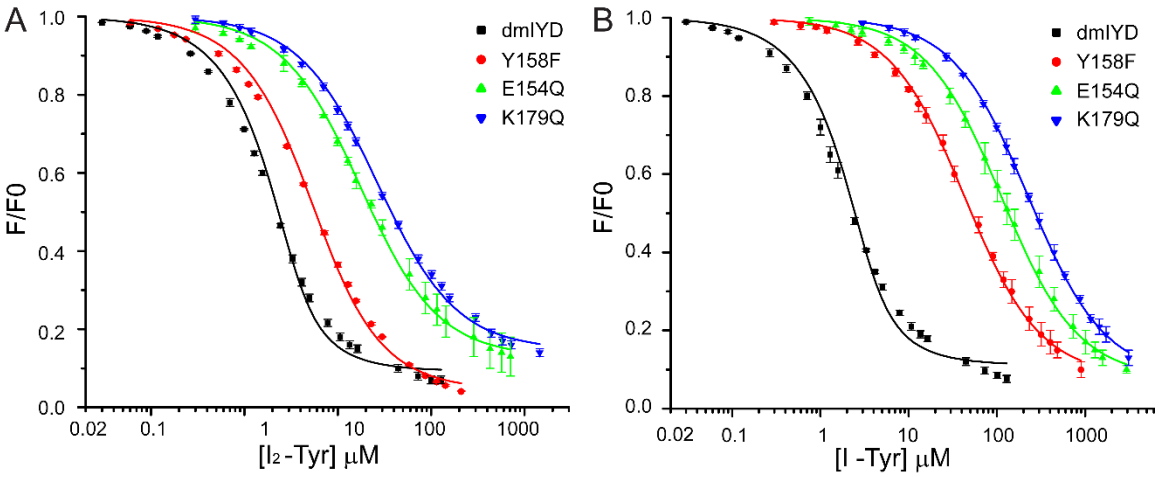

Figure 3-5. Quenching of FMN fluorescence is indicative of I_2-Tyr (A) and I-Tyr (B) binding in the active site of dmIYD and its mutants. Each data point represents an average of 3 individual observations and error bars represent standard deviation from the averages. The K_d values are derived from the best fit of the data (solid lines) to equation 3.

Table 3-2. Impact on substrate coordination after active site mutagenesis of dmIYD.

	K_d (µM)	
Enzyme	**I_2-Tyr**	**I-Tyr**
dmIYD[a]	0.54 ± 0.05	0.62 ± 0.08
dmIYD Y158F[a]	3.8 ± 0.2	42 ± 2
dmIYD E154Q[a]	16 ± 0.8	112 ± 5
dmIYD K179Q[a]	25 ± 1	230 ± 10
mmIYD[b]	-	2 ± 0.2
mmIYD Y157F[b]	-	40 ± 10
mmIYD E153Q[b]	-	>1000

[a]Determined from data illustrated in Fig. 3-5. Errors derive from least square fitting. [b]Parameters were previously reported.[29]

3.3.4 Contribution of zwitterion coordination residues towards catalysis

The catalytic efficiency of dmIYD for turnover of [^{125}I]-I_2-Tyr was within the range of the catalytic efficiencies of previously assayed IYD homologs (Table 2-3 and Table 3-3).

The Y158F mutation resulted in an 8-fold increase in K_m while k_{cat} was increased by little over 2-fold (Table 3-3). The overall impact on k_{cat}/K_m was a modest 3.5-fold decrease compared to dmIYD. The K179Q mutation resulted in a 64-fold increase in K_m while k_{cat} was not affected significantly. The catalytic efficiency was lowered by 55-fold compared to dmIYD. The E154Q mutation had a large effect on K_m which increased over 300-fold while the k_{cat} was modestly reduced by nearly 3-fold. The catalytic efficiency was most adversely affected with an 860-fold decrease compared to dmIYD.

Table 3-3. Impact of mutating substrate zwitterion coordination residues on deiodination of I_2-Tyr.

Enzyme	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m (min^{-1} μM^{-1})
dmIYD[a]	11 ± 2	6.5 ± 0.4	0.6 ± 0.1
dmIYD Y158F[a]	90 ± 20	15 ± 2	0.17 ± 0.04
dmIYD K179Q[a]	700 ± 100	7.7 ± 0.5	0.011 ± 0.002
dmIYD E154Q[a]	3400 ± 500	2.4 ± 0.2	0.0007 ± 0.0001
mmIYD[b]	65 ± 16	9.3 ± 1.6	0.23 ± 0.05
mmIYD Y157F[b]	440 ± 170	65 ± 16	0.15 ± 0.07

[a]Determined from data illustrated in Appendix Fig. B8. Errors derive from least square fitting. [b]Parameters were previously reported.[29]

3.4 Discussion

3.4.1 Possible substrates for dmIYD

Halotyrosines I-, Br- and Cl-Tyr are all substrates for dmIYD in vitro. The enzyme did not exhibit substantial preference for any of these substrates based on nearly identical binding affinities and similar catalytic efficiencies of dehalogenation. The k_{cat} values correlate to the phenyl carbon-halogen (C-X) bond strength with the highest k_{cat} measured for the weakest C-I bond and lowest k_{cat} measured for stronger C-Cl bond. The differences in k_{cat} are small compared to the differences in C-X bond strength that increases by about

15 kcal mol^{-1} each from C-I to C-Br and C-Br to C-Cl.[67] The k_{cat}/K_m for steady state turnover of I-Tyr and Br-Tyr by dmIYD (0.5 min^{-1} μM^{-1}) is similar to the second order pre-steady state rate constants (k_{ox}) for single turnover of I-Tyr and Br-Tyr by human (hs) IYD (0.52 min^{-1} μM^{-1} and 0.44 min^{-1} μM^{-1}).[62] However, the k_{cat}/K_m for steady state turnover of Cl-Tyr (0.14 min^{-1} μM^{-1}) by dmIYD is about 5- to 6-fold higher than k_{ox} for hsIYD (0.024 min^{-1} μM^{-1}) suggesting that Cl-Tyr is processed more rapidly by dmIYD than hsIYD. Consistent with earlier observations involving mammalian IYDs,[19,31] F-Tyr binds relatively weakly to dmIYD and generates no detectable product in presence of 25-fold higher enzyme concentration and 6-fold longer incubation time ($k_{cat} < 0.002$ min^{-1}) than used for other halotyrosines precludes it as a substrate. This is not surprising since fluorinated tyrosine residues are not known to occur naturally[68] and these enzymes have consequently not evolved to reduce the strong C-F bond.

Little information exists about function of iodine in insects. Unlike chordates, insects are not known to synthesize THs which require iodine. Other proteins involved in TH synthesis and metabolism and iodide uptake including TPO, IDs and the NIS do not have reported homologous counterparts in *Drosophila*.[57] Very early studies involving feeding radioactive ^{131}I isotope to *Drosophila gibberosa* larvae had indicated iodide accumulation in specific areas of the larval cuticle especially in posterior ventral regions and near air passages and mouth regions suggesting incorporation via specific metabolic processes.[69] It is possible that iodotyrosines are produced by action of peroxidase enzymes during insect cuticle sclerotization and are substrates for dmIYD. Br-Tyr is produced by hypobromous acid generated through eosinophil peroxidase activity during inflammatory responses in mammals and has recently been observed as a substrate for IYD.[70] Bromine has been

demonstrated to be essential for collagen crosslinking via hypobromous acid produced by peroxidasin activity and the complete absence of bromine in food sources can be lethal in *Drosophila*.[71] Hypobromous acid may also brominate tyrosine residues and produce bromotyrosine for later processing by dmIYD. Chlorotyrosines have been reported in the sclerotized cuticle of locusts and other insects.[72] It is possible that dmIYD plays a catabolic role by degrading the halogenated tyrosine residues produced as by-products of other metabolic reactions or dmIYD could be involved in processing of halotyrosines in a yet undiscovered function in insects and other lower invertebrates.

3.4.2 Relative Importance of Active Site Residues

The three active site residues Glu, Tyr and Lys coordinating the halotyrosine substrate zwitterion via their side chains are conserved in IYDs from mammals to prokaryotes investigated to date and are thought to be general markers for halotyrosine substrate recognition within the IYD subgroup of the nitro-FMN reductase superfamily. Previously, active site mutations had been attempted with mmIYD, but additional non-native mutations at cysteine residues (Cys to Ala) were necessary for bacterial expression of soluble protein. While limited information could be gathered for these mutants (Table 3-2 and Table 3-3), the catalytic efficiency could not be accurately quantified for the E153Q mutant. Furthermore, the K178Q mutant was insoluble and precluded further analysis.[29] In contrast, all 3 mutations were well tolerated at equivalent residues of dmIYD and this facilitated the determination of their relative impact on binding affinity as well as catalysis. The active site E154 is most important for efficient catalysis and a conservative change of E154Q substituting the negative carboxylate for amide resulted in a decrease in catalytic efficiency by nearly 3 orders of magnitude. However, this loss in catalytic efficiency cannot be

entirely attributed to the loss in binding affinity as previously thought for the equivalent E153Q mutant of mmIYD[29], since an alternative K179Q mutant of dmIYD displayed greater loss in binding affinity for I_2-Tyr than the E154Q mutant (46- versus 30-fold) but maintained significantly higher (16-fold) catalytic efficiency. Both ligand affinity and catalysis are least affected by the Y158F mutation that eliminates hydrogen bonding between a phenolic OH and the substrate zwitterion. The overall impact of active site mutations on catalysis are likely linked to the ability of the enzyme-substrate complex to re-position an active site Thr to enable hydrogen bonding to the FMN N^5. The importance of such a substrate-dependent activation of the FMN cofactor was indicated from previous structural and catalytic studies of hsIYD as well as catalytic studies of prokaryotic hhIYD.[31,73]

Previous studies on hsIYD indicated preferential binding of the deprotonated phenolate form of the substrate compared to its neutral phenol form.[31] Coordination of the substrate phenolate to both the 2'-hydroxy group of FMN and a backbone amide hydrogen as observed in all co-crystal structures of IYD·I-Tyr supports this observation.[31,28] However, the K_d values of I-Tyr and I_2-Tyr are not indicative of the greater proportion of the phenolate form of I_2-Tyr (91%, pK_a 6.4)[74] versus I-Tyr (7%, pK_a 8.5)[75] under assay conditions (pH 7.4). Instead for hsIYD, the K_d value for I-Tyr is smaller than I_2-Tyr.[31] Accommodation of the added steric bulk of an additional iodo substituent requires some structural changes in the active site that likely dominate the higher affinity for the predominantly phenolate form of I_2-Tyr. In the case of dmIYD, these opposing properties may balance each other since the K_d values for I-Tyr and I_2-Tyr are nearly indistinguishable (Table 2). This balance is not maintained in the three active site mutants. Even though the

mutated residues most directly affect the coordination of the zwitterion that is common to both I-Tyr and I_2-Tyr, their affinity for I-Tyr is 7- to 10-fold lower than for I_2-Tyr (Table 3-2). This is the first time the bulkier I_2-Tyr substrate has exhibited a lower K_d value and may reflect a greater sensitivity to the significantly higher fractional population of the phenolate form in solution. However further evidence is required though pH profile studies on dmIYD and its mutants.

Overall, the relative impact of active site residues on substrate recognition and catalysis should be broadly applicable to IYD homologs within the nitro-FMN reductase superfamily given the high degree of conservation of these residues from archaea to mammals. Our findings should facilitate the annotation of enzymes within the nitro-FMN reductase superfamily with the active site Glu being the most robust indicator of a halotyrosine dehalogenase followed by Lys while Tyr is more replaceable than other active site residues. We identified archaeal nitro-FMN reductase superfamily members in which the active site Tyr is replaced with His but the requisite Glu and Lys are conserved such as putative NADH oxidases from *Hyperthermus butylicus* (Accession WP_011822328.1) and *Staphylothermus marinus* (Accession WP_011839241.1). These proteins are likely to function as halotyrosine dehalogenases.

3.5 Summary

Efficient and comparable dehalogenation of I-Tyr, Br-Tyr as well as Cl-Tyr is demonstrated by dmIYD. These results indicate that the enzyme may function more broadly as a halotyrosine dehalogenase rather than just an iodotyrosine deiodinase, an annotation ascribed to it following the discovery of its primary function of iodine salvage

in the mammalian thyroid. Through site directed mutagenesis of active site residues, we demonstrated that Glu coordinating the substrate zwitterion is most important for catalysis followed by Lys. The Tyr residue has the smallest effect on the catalytic efficiency. The impact of mutating these active site residues on catalysis is not entirely due to loss in binding affinity as previously thought since the K179Q mutant actually displays weaker affinity for ligand than the E154Q mutant. Our results should further facilitate annotation of halotyrosine dehalogenases within the nitro-FMN reductase superfamily of enzymes.

Chapter 4: Silencing the halotyrosine dehalogenase gene in *Drosophila* using RNAi

4.1 Introduction

The fruitfly has been studied as a genetic model organism for over a century.[76,77] Its genome encodes approximately 15,000 genes and has more homologs to human genes than *C. elegans*, a worm model organism.[78] Traditional genetic studies relied on abnormal phenotypes to discover gene function, an approach known as forward genetics. Since the *Drosophila melanogaster* (referred to as simply *Drosophila* hence forth) genome was sequenced in 2000,[79] geneticists began to utilize the strategy entitled reverse genetics that involves modifying a known gene sequence and investigating its phenotypic manifestations.[80] *Drosophila*, being one of the most genetically tractable organisms enabled rapid development of tools for genetic investigations such as balancer chromosomes with multiple inversions (segments of chromosomes with reversed orientations), dominant phenotypic markers and recessive lethal mutations.[81] Recombination between homologous chromosomes with and without inversions is suppressed during meiosis.[82] *Drosophila* strains with a chromosome carrying a lethal gene mutation can be maintained through pairing with a corresponding balancer chromosome and obviating the need to perform genetic screens for each generation.[83]

A widely used reverse genetics approach relies on gene silencing by RNA interference wherein the mRNA for a particular gene is degraded after recognition by double stranded RNA of complementary sequence.[84] A heterologous yeast expression system employing the GAL4 transcriptional activator recognizing the upstream activating sequence (UAS) has been widely employed to drive expression of RNAi in *Drosophila*.[85,86] GAL4

expression can be controlled by tissue and cell type specific *Drosophila* promoters enables targeted silencing of a gene of interest. Each component of this dual system (GAL4 and UAS) is maintained in separate parent strains (usually balanced) and is inherited by the offspring leading to silencing of the target gene when the selected promoter is endogenously expressed. GAL4 allows for modulation of gene silencing in a temperature dependent manner (optimal activity at 29 °C and minimum activity at 16-18 °C in *Drosophila*).[85]

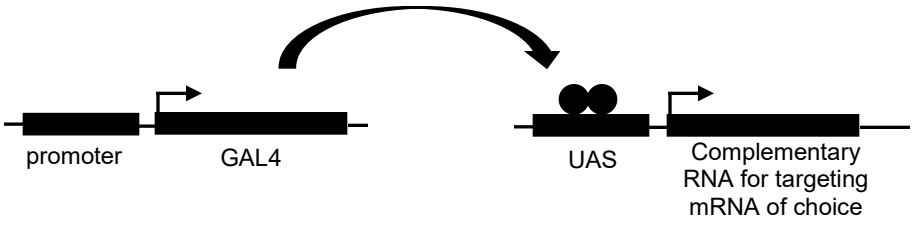

Figure 4-1. The GAL4-UAS system commonly used for gene silencing in *Drosophila*. GAL4 expression is controlled by a chosen species specific promoter. This figure is adapted from Brand and Perrimon.[85]

Previous genome wide profiling of the *Drosophila* transcriptome indicated high expression of the halotyrosine dehalogenase encoding gene (CG6279) in the adult male testes with moderate expression in the pre-adult stages (Fig. 4-2).[87,88]

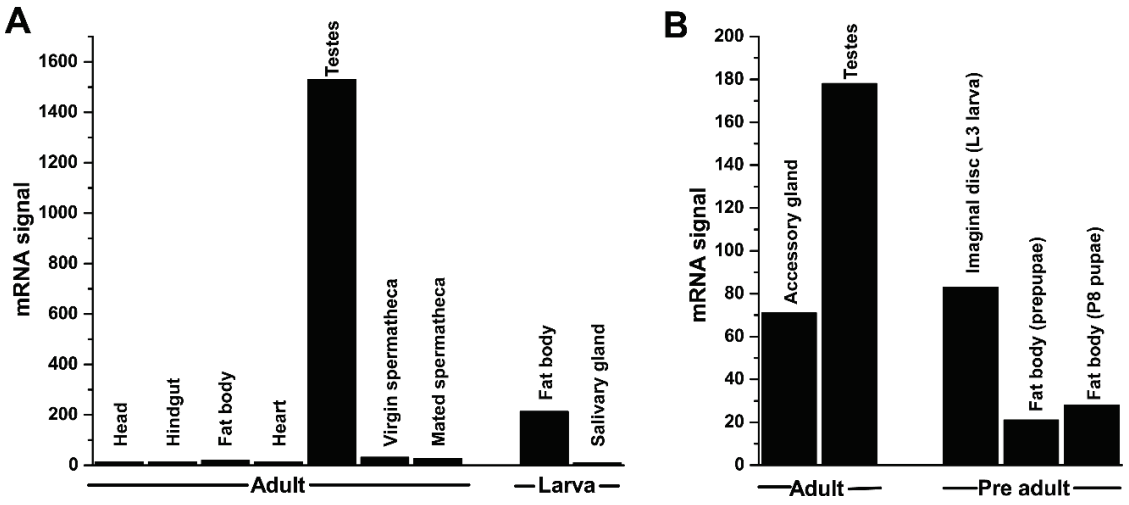

Figure 4-2. Reported mRNA expression of IYD in different *Drosophila* tissues. (A) Expression levels reported in the FlyAtlas study (microarrays).[87] B. Expression levels reported by the modENCODE consortium (RNA-Seq).[88]

The high levels of CG6279 mRNA in the testes suggested a distinct role for the dehalogenase in spermatogenesis. In *Drosophila* testes, spermatogenesis begins with 4 rounds of mitotic division of the germline to produce a 16 cell spermatogonia which then transition into spermatocytes and wherein meiosis is initiated (Fig. 4-3).[89] After 2 rounds of meiosis, the 64 cell spermatids undergo elongation and individualization during the late stages of spermatogenesis.[90] The developing germline is surrounded by somatic cyst cells throughout their differentiation process and are known to regulate sperm development.[91] If the dehalogenase is indeed involved in spermatogenesis, gene silencing should affect the fertility of the males and result in a testes specific phenotype.

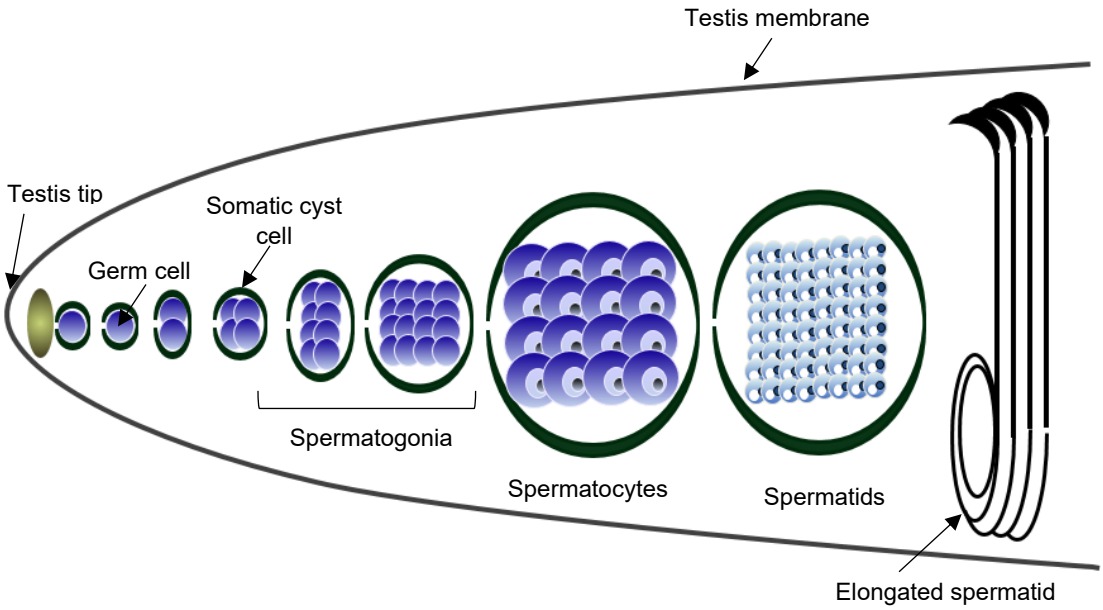

Figure 4-3. Germline lineage in *Drosophila* spermatogenesis. Development proceeds from the left (tip of testes) to the right as indicated in figure. Figure was provided by Dr S. Eun.[89]

In this chapter we investigate the expression pattern of CG6279 mRNA in *Drosophila* testes through *in situ* hybridization. Using the reverse genetics approach we also investigate the effects of CG6279 gene silencing on male fertility and testes morphology. All experiments described in this chapter were performed in collaboration with Dr. S. Eun (X. Chen Laboratory, JHU Biology).

4.2 Experimental Procedures

4.2.1 *Drosophila* Stocks

Drosophila were raised on standard yeast, molasses, cornmeal and agar medium and maintained at 25 °C. Yellow White (*yw*) reference *Drosophila* strain and appropriately balanced GAL4 expression strains including *nos-GAL4* (3rd chromosome); *bam-GAL4, UAS-GAL4* (2nd chromosome); *C587-GAL4* (X chromosome); *tj-GAL4* (2nd chromosome),

eya-GAL4 (2nd chromosome), *ubi-GAL4* (2nd chromosome) and *UAS-dcr2* were used from stocks available in the laboratory of Dr X. Chen (Dept. of Biology, JHU). The RNAi expression strains *UAS-37267 RNAi* (3rd chromosome) and *UAS-105378 RNAi* (2nd chromosome) were obtained from the Vienna Drosophila RNAi center (VDRC, Vienna, Austria). Balanced *UAS-shmiRNA-A*, *UAS-shmiRNA-AB* and *UAS-shmiRNA-B* stocks were generated as described below. Drosophila from selected GAL4 expressing strains were mated with chosen UAS-RNAi strains to obtain progeny expressing interfering RNA.

4.2.2 In situ hybridization

RNA i*n situ* hybridization (ISH) for detection of CG6279 transcript was performed as previously described.[92,93] Expression pattern of CG6279 was detected using whole mount testes from the *yw* strain. Primers 5'-CATGAAGTTTCGGTAGAAGAG–3' (forward) and 5'– CATCTGCACTTGCTGGCTATT–3' were used to amplify a 405 bp region of isoform A transcript of CG6279 (CG6279-RA) from a *Drosophila* testes cDNA library (kindly provided by Dr S. Eun, Chen lab, Dept. of Biology, JHU) by using *Taq* DNA polymerase (Thermo Scientific). Similarly, primers 5'-ATTGTGGAACAGGAGGAGCTG-3' (forward) and 5'-ATTCTTTCTCGCCAAGTCGGG-3' (reverse) were used to amplify a 393 bp region of CG6279 that is common to both isoforms A and B of CG6279 (CG6279-RA and CG6279-RB). The PCR products with 3'-A overhangs (added by Taq DNA polymerase in a non-template dependent manner) were then ligated into the linearized pGEM-T Easy plasmid (Promega) with 3'-T overhangs using T4 DNA ligase (Promega) and transformed into GeneHogs *E. coli* cells (Invitrogen). Transformants containing inserts (white colony) were identified via blue-white colony screening from agar plates coated with IPTG and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal). Plasmids were

purified from single colonies and the desired inserts were subsequently subcloned into the pBluescript II vector using SacII and PstI restriction sites. Transformants containing inserts were identified via blue-white colony screening and the orientation of the inserts was determined through DNA sequencing (Genewiz). Based on the orientation of the DNA inserts, anti-sense and sense probes were synthesized using T3 and T7 RNA polymerases (Roche Life Science) with PstI or SacII linearized vector as template respectively. A Digoxigenin (DIG) label was incorporated in the probes through DIG labeled UTPs (DIG RNA labeling mix, Roche Life Science). All procedures involved use of RNAase free water (Sigma).

All subsequent procedures including hydrolysis of RNA probes, testes preparation and in situ hybridization were performed as previously described.[92,93] Imaging of the testes was performed using a Zeiss HXP120C Apotome microscope.

4.2.3 Fertility testing

Newly eclosed single males of desired genotype were mated with two virgin *yw* females each in vials with standard yeast, molasses, cornmeal and agar medium. Mating was allowed for 7 days at 25 °C after which adults were removed and vials were evaluated for fertility and fecundity (number of progeny produced by each male). The progeny emerging from each vial were quantified from day 10 to day 18. If any of the adults died during the mating period, the vial was excluded from analysis.

4.2.4 Design of UAS-shmiRNA constructs for transgenic *Drosophila*

Guidance for generating transgenic RNAi *Drosophila* strains was provided by Dr. Shekerah Primus (Van Doren Lab, Dept. of Biology, JHU). The open reading frame of CG6279-RA was analyzed for potential siRNA sites (21nt) by using the DSIR program

(Design of Small Interfering RNA).[94] Sequences targeting CG6279-RA, both CG6279-RA and CG6279-RB and 5'-untranslated region (UTR) of CG6279-RB were selected (Table 4-1). Scanning for potential off-target sites was performed as previously described.[95]

Table 4-1. Chosen siRNA sequences using DSIR

Isoform	Sequence (5'-3')
CG6279-RA specific	
Sense	GGAAGUGUGGAGAGAUAAUGA
Antisense	AUUAUCUCUCCACACUUCCUU
CG6279-RA and RB	
Sense	GCAUUACUACAACGAGAUAUC
Antisense	UAUCUCGUUGUAGUAAUGCCU
CG6279-RB 5' UTR	
Sense (R)[a]	GCAGUUACAUCUUGAAGCACC
Antisense (R)[a]	UGCUUCAAGAUGUAACUGCUU

Cloning of the DNA sequences into the Valium 20 vector for generating short hairpin micro RNA (shmiRNA) constructs followed a protocol provided by the Transgenic RNAi Project (TRiP).[96] Incorporation of insert was confirmed by DNA sequencing (Genewiz.) Vectors with the desired inserts were purified by midi-prep (GeneJet plasmid midiprep kit, Thermo Scientific) and 40 µg of each vector (1µg/µl) was shipped to Bestgene Inc (Chino Hills, CA) for injection into embryos from strain attP40 containing an attP recombination site on the 2nd chromosome to generate *UAS-shmiRNA-A*, *UAS-shmiRNA-AB* and *UAS- shmiRNA-B* strains. Vermillion^{+} screening of *Drosophila* transformants was provided by Bestgene Inc. Presence of the insert for shmiRNA within the genome was confirmed by PCR and agarose gel electrophoresis from single males using the forward primer 5'-

ACCAGCAACCAAGTAAATCAAC-3' and reverse primer 5'-TAATCGTGTGTGATG-CCTACC-3'.

4.2.5 Microscopic analysis of *Drosophila* testes

Males were dissected to isolate intact testes using a dissection microscope. The testes were mounted on a microscopic slide using phosphate buffered saline. The morphology of the testes was examined via phase contrast microscopy using a Zeiss HXP120C Apotome microscope. At least 10 pairs of testes for each male genotype were evaluated by microscopy.

4.3 Results

4.3.1 In situ hybridization for detection of CG6279 transcription

In order to probe the region of the *Drosophila* testes that shows presence of CG6279 mRNA, we designed anti-sense RNA probes. One of the probes specifically targeted the N-terminal domain of CG6279-RA and another targeted the deiodinase domain common to both CG6279-RA and CG6279-RB. The testes were analyzed for color development using an alkaline phosphatase (AP) conjugated antibody targeting the DIG labeled RNA probes and AP substrate 5-bromo-4-chloro-3-indolyl phosphate with nitro blue tetrazolium salt. A strong signal was observed in the region occupied by post meiotic germ cells mainly spermatocytes, spermatids and elongated spermatids but not closer to the tip where germ cells in early stages of spermatogenesis are housed (Fig. 4-4A and B) as indicated in the schematic representation depicted in Fig. 4-3.

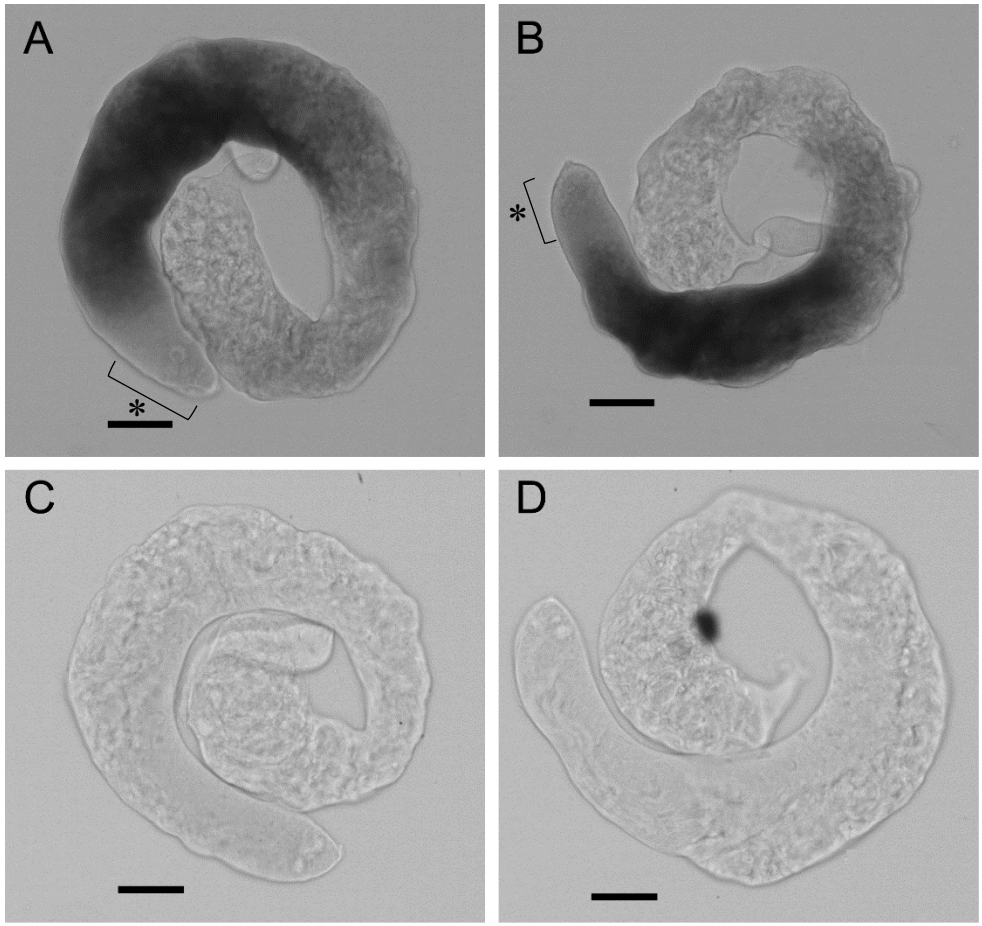

Figure 4-4. In situ hybridization of *Drosophila* testes with DIG labeled RNA probes. Anti-sense probes targeting CG6279-RA only (A) and both CG6279-RA and RB (B) show intense staining in the region of the testes housing post meiotic germ cells. Region of testes normally occupied by early germ cells undergoing mitotic division is indicated by (*).[97,98] Corresponding sense RNA probes (C and D) were used as controls. Scale bars represent 50 μm.

The region of testes that are stained and their intensity of staining is similar for the antisense RNA probe targeting CG6279-RA only as well as the probe targeting both isoforms (CG6279-RA and CG6279-RB). Sense RNA probes (controls) did not show any staining pattern (Fig. 4-4 C and D).

4.3.2 Fertility and testes phenotype of UAS-RNAi strains from VDRC

UAS-RNAi strains from VDRC targeting CG6279 transcripts were initially tested in combination with *bam-GAL4* driving expression specifically in germ cells and *eya-GAL4*

and *c587-GAL4* driving expression in somatic cells. *UAS-105378 RNAi* targets CG6279-RA while *UAS-37267 RNAi* targets both CG6279-RA and CG6279-RB (Appendix Fig. C1). Males of the genotype *eya-GAL4*/*UAS-105378 RNAi* displayed 50% sterility (7 of 14 *eya-GAL4*/*UAS-105378 RNAi* males tested were sterile while all control males were fertile) and a 92% decrease in fecundity (average number of progeny per *eya-GAL4*/*UAS-105378 RNAi* male was 11 compared to 141 per control male). Other GAL4-UAS-RNAi systems did not show significant difference compared to controls (Fig. 4-5).

Males of the genotype *eya-GAL4/UAS105378 RNAi* were dissected and their testes were analyzed by phase-contrast microscopy. A phenotype was observed showing aggregation of cells during the late stages of spermatogenesis (Fig. 4-6A) with very few or no motile sperm. In contrast, the control (*UAS-105378 RNAi*/+) showed normal morphology (Fig. 4-6B) and large number of motile sperm.

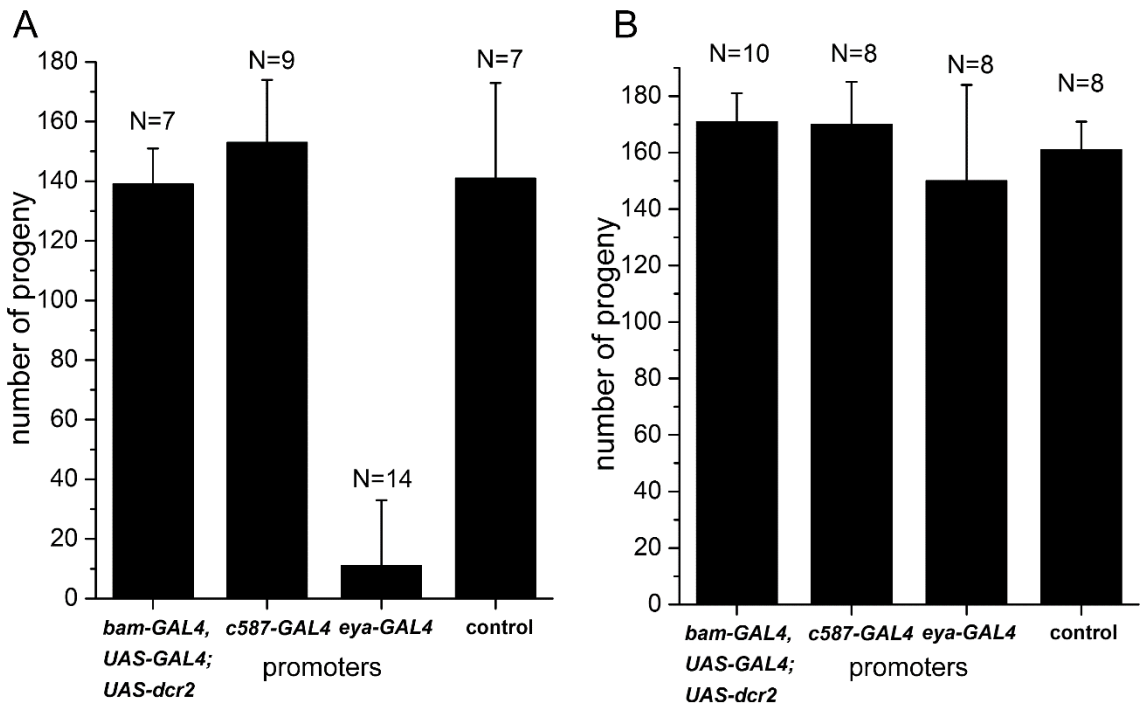

Figure 4-5. Fertility of male *Drosophila* expressing *UAS-105378 RNAi* (A) and *UAS-37267 RNAi* (B) driven by GAL4 under control of the indicated promoter. N represents the number of individual males tested. Error bars represents standard deviation from average number of progeny produced. Control males are of genotype +/*UAS-RNAi*.

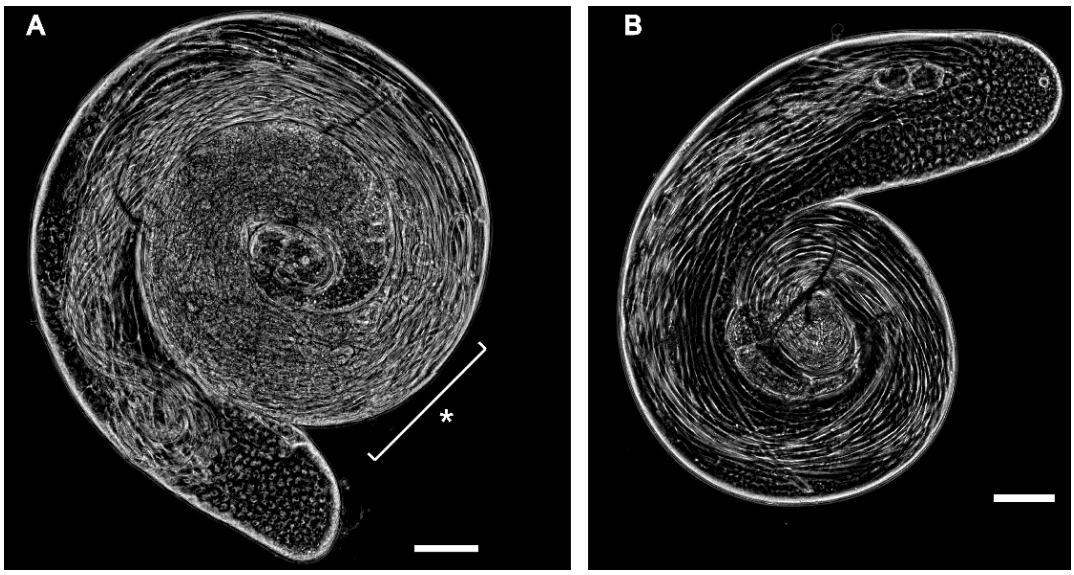

Figure 4-6. Phase contrast microscopy images of *Drosophila* testes. (A) Testis from *eya-GAL4/UAS-105378 RNAi* males. Late stage cell aggregation phenotype is indicated with (*). (B) Control from +/*UAS-105378 RNAi* male. Males were maintained at 25 °C for 3 to 4 days after eclosion prior to dissection. Scale bar represents 100 μM.

4.3.3 Design of *UAS-shmiRNA* transgenic *Drosophila* strains and evaluation of testes phenotype

The transgenic *Drosophila* strains from VDRC generate long dsRNA from inverted repeat inserts (300-500 bp) under control of a UAS promoter. The dsRNA is further processed by dicer to generate siRNAs.[99] In order to selectively target the CG6279-RB encoding only the deiodinase domain of CG6279, we designed a short hairpin micro-RNA targeting the 5'UTR of CG6279-RB (*UAS-shmiRNA-B*) which is not common to CG6279-RA. We also designed *UAS-shmiRNA-A* and *UAS-shmiRNA-AB* that target the isoform A transcript and the common region of both transcripts respectively (Appendix Fig. C2) to further verify observations with the VDRC strains.

We evaluated *eya-GAL4/UAS-shmiRNA* males for a testes specific phenotype. All *eya-GAL4/UAS-shmiRNA* strains were maintained at 29 °C as opposed to 25 °C to maximize GAL4 activity. A late stage cell aggregation phenotype with few or no motile sperm was observed with all 3 *UAS-shmiRNAs* (Fig. 4-7A to C). This phenotype was similar to previously observed phenotype in tested of eya-*GAL4/UAS105378 RNAi* males. However *eya-GAL4/+* control males not encoding UAS-RNAi also showed a similar late stage phenotype (Fig. 4-7D). Additional germ cell specific (*nos-GAL4*), somatic cell specific (*tj-GAL4*) and general GAL4 (*ubi-GAL4*) drivers failed to produce any testes phenotype in combination with UAS-RNAi at 29 °C.

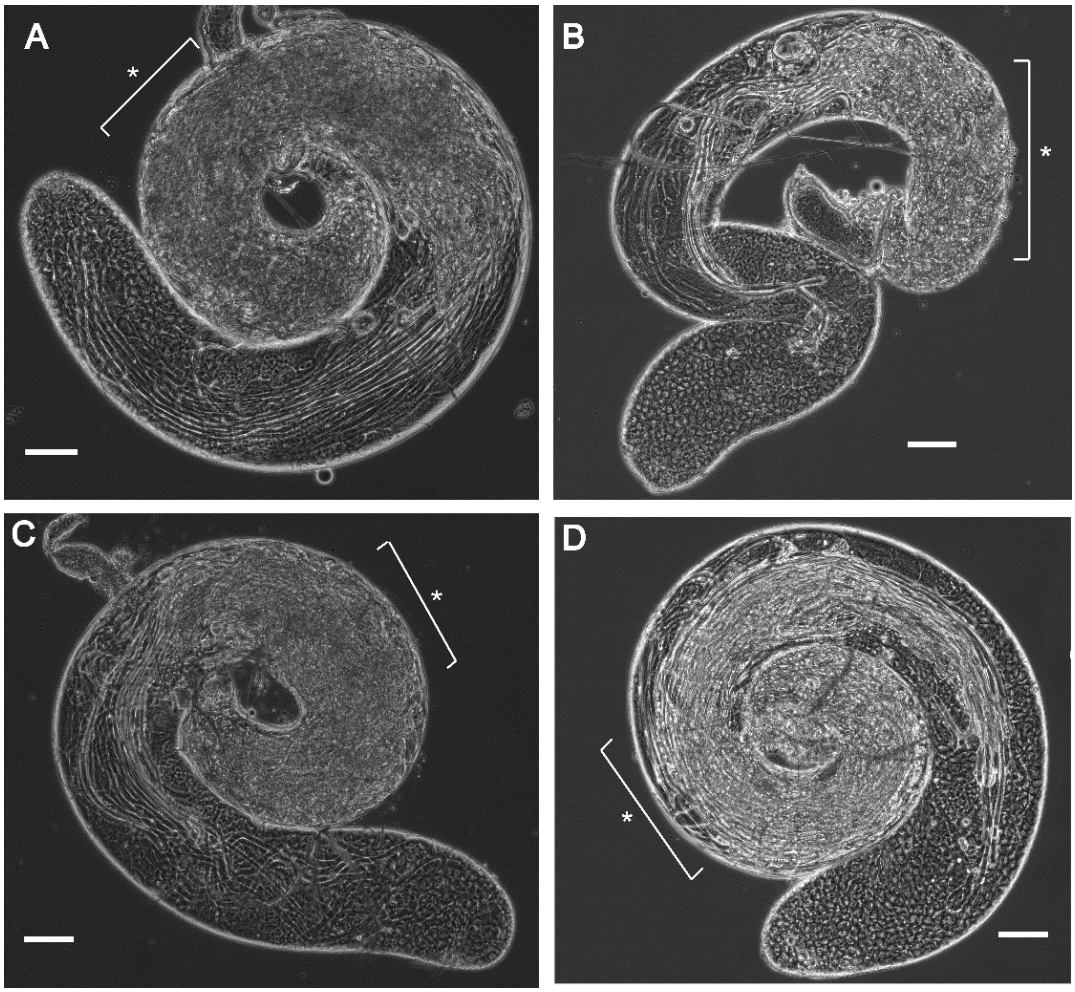

Figure 4-7. Phase contrast microscopy images of *Drosophila* testes from *eya-GAL4/UAS-shmiRNA-A* (A), *eya-GAL4/UAS-shmiRNA-B* (B), *eya-GAL4/UAS-shmiRNA-AB* (C) and *eya-GAL4/+* control males (D). Late stage cell aggregation phenotype is indicated with (*). Males were maintained at 29 °C for 3-4 days prior to dissection. Scale bar represents 100 μM.

4.4 Discussion

High expression of CG6279 in the testes of *Drosophila* males reported previously published studies[87, 88] was confirmed through in situ hybridization experiments using anti-sense RNA probes targeting the CG6279 mRNA. The staining pattern revealed high expression of the dehalogenase gene in late stages of spermatogenesis including regions housing elongated spermatids. RNA probe targeting CG6279-RA that also includes a large

N-terminal domain and the RNA probe targeting the deiodinase domain of both CG6279-RA and CG6279-RB show similar staining patterns and signal intensity.

High expression of the dehalogenase gene in the testes is suggestive of an important function in spermatogenesis. We initially investigated the effect of suppressing the CG6279 transcript in cell-type specific (germ cell and somatic cell) experiments on male fertility with RNAi transgenic *Drosophila* from VDRC targeting CG6279. Males expressing RNAi targeting CG6279-RA using a somatic cell specific promoter 'eya' displayed over 10-fold lower fecundity (50% of males tested were sterile) compared to controls lacking the *eya-GAL4* driver. Germ cell specific RNAi expression did not affect fertility or fecundity. This indicated that the dehalogenase gene could have a role in somatic cyst cells. Microscopic evaluation of the testes from *eya-GAL4/UAS-105378 RNAi* males showed a late stage cell aggregation phenotype. This phenotype was observed in the region of the testes just following the region in which CG6279 is transcribed in high levels (Figs. 4-4 and 4-6).

Previous studies have indicated more efficient gene silencing by shmiRNA compared to siRNA derived from long dsRNA.[100] In addition, RNAi through expression of long dsRNAs has been shown to result in off-target gene silencing.[95] Therefore to confirm our preliminary observations and target each transcript isoform selectively, additional transgenic *Drosophila* strains encoding shmiRNA targeting CG6279 transcripts were generated. *UAS-shmiRNA-A*, *UAS-shmiRNA-B* and *UAS-shmiRNA-AB* expression driven by *eya-Gal4* produced the same male testes phenotype as previously observed with *UAS-105378 RNAi* expression driven by *eya-GAL4*. However, *eya-GAL4/+* control males that do not encode any RNAi showed the same phenotype suggesting that the original

phenotype observed was likely a false positive and did not result from CG6279 gene silencing but was related to the *eya-GAL4* transgene itself.

We were unable to reproduce the phenotype observed for *eya-GAL4* driven RNAi expression using any other germ cell or somatic cell specific GAL4 driver and UAS-RNAi combination including a general strong driver like *ubi-GAL4*.

In conclusion, we were unable to determine the effect of CG6279 gene silencing on *Drosophila* spermatogenesis using RNA interference. It is possible that gene silencing was not efficient for CG6279 using the *GAL4–UAS-RNAi* system and the observed phenotype was an artifact of the *eya-GAL4* transgene. Editing of CG6279 at the genomic level to generate CG6279 loss of function *Drosophila* strains would enable investigation of the biological role of G6279 most definitively.

4.5 Summary

The expression of the CG6279 gene encoding *Drosophila* halotyrosine dehalogenase in the testes was confirmed by in situ hybridization using antisense RNA probes. Gene silencing via RNA interference targeting CG6279-RA driven by somatic cell specific *eya-GAL4* resulted in lower fertility and fecundity of males due a late stage cell aggregation testes phenotype with few or no motile sperm. However this phenotype was also observed in control male testes devoid of *UAS-RNAi* but encoding the *eya-GAL4* driver system. The phenotype was not reproduced using any other GAL4 driver. Therefore the effect of CG6279 gene silencing using RNA interference on *Drosophila* testes was inconclusive and will require an alternative strategy for generating dehalogenase deficient transgenic *Drosophila*.

Chapter 5: CRISPR/Cas9-based editing of a *Drosophila* dehalogenase gene

5.1 Introduction

The gene encoding the dehalogenase enzyme (CG6279) is highly expressed in *Drosophila* testes in the late stages of spermatogenesis. RNAi experiments aimed at CG6279 silencing in the testes were inconclusive. An alternative approach to affect dehalogenase expression would eliminate the need for targeting the mRNA through RNAi. In the past few years, a RNA guided (CRISPR) DNA nuclease (Cas9) system has been widely adopted for genome editing due to its simplicity of design.[101]

Clustered regularly interspaced short palindromic repeats or CRISPR are DNA repeats separated by spacer sequences found in prokaryotic genomes and are a component of prokaryotic defense mechanisms against invading phages and viruses.[102] CRISPR RNA delivers an associated endonuclease to complementary DNA through base pairing. The DNA target is then cleaved by the endonuclease resulting in double stranded breaks (DSB).[103] The CRISPR-associated endonuclease (Cas9) from *S. pyrogenes* has been shown to require a single guide RNA (gRNA) recognizing a 20 nucleotide DNA sequence with an NGG sequence (protospacer adjacent motif or PAM) immediately adjacent to it. Cas9 cleaves the target DNA three nucleotides upstream of PAM (Fig. 5-1).[104] Earlier techniques for genome editing involved programmable nucleases such as zinc finger nucleases (ZFNs)[105] and transcription activator-like effector nucleases (TALENS)[106] but these required protein engineering for each target sequence recognition. Engineering these nucleases in the genomic context for multiple targets is challenging, laborious and requires extensive validation for each individual target sequence.[103,107] CRISPR/Cas9 allows precise targeting of any 20 nt DNA sequence adjacent to a PAM through routine

replacement of gRNA. This ease of design has led to the rapid replacement of older genome editing technologies with CRISPR/Cas9.

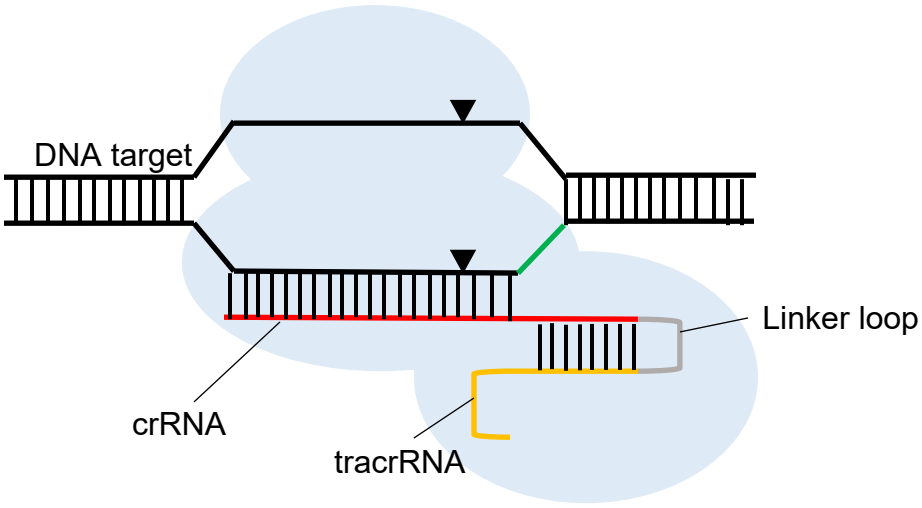

Figure 5-1. DNA recognition by CRISPR/Cas9. Cas9 (blue) is guided by a single chimera between the RNA component targeting the DNA sequence of interest (crRNA, red) and the RNA associating with Cas9 (tracrRNA, yellow) to form the guide RNA (gRNA). PAM is indicated in green. Cleavage sites are indicated by wedges. Figure adapted from Jinek et al.[104]

CRISPR/Cas9 has been effectively used in *Drosophila* to disrupt gene function by inducing double stranded breaks at targeted genes that are subsequently repaired by non-homologous end joining to produce small insertions and deletions at target sites causing frame-shift, missense or nonsense mutations.[108-111] Using 2 gRNAs targeting the 5'end and 3'end of a fruitfly gene locus, defined deletions have been introduced to generate null mutants.[108,110] Homology directed repair (HDR) has been employed for precise gene editing through CRISPR/Cas9 in *Drosophila*.[110,112] For HDR, a DNA repair template with flanking homology arms upstream and downstream of the DSB and the desired insert or mutation in between the flanking sequence is co-injected with CRISPR components (Fig.

5-2). Both single strand oligodeoxynucleotide (ssODN) repair templates (less than 200 bp) as well as long double stranded DNA for incorporation of large exogenous sequences including screenable markers have been successfully employed in *Drosophila*. Transgenic fly strains constitutively expressing the Cas9 enzyme in the germline (*nos-Cas9* and *vasa-Cas9*) are available for engineering inheritable mutations through germline engineering.[108,112]

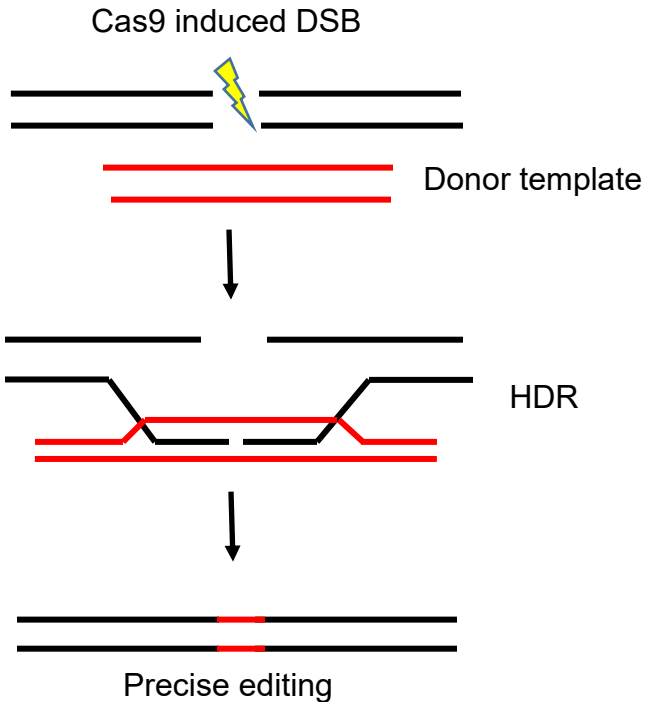

Figure 5-2. Homology directed repair (HDR) following double stranded break (DSB) by Cas9. Figure is adapted from Sander and Joung.[103]

In this chapter we describe editing the CG6279 gene encoding the dehalogenase enzyme to generate loss of function mutant fruitfly strains followed by testing males carrying these mutations for fertility to assess whether the halotyrosine dehalogenase is essential for spermatogenesis.

5.2 Experimental Procedures

5.2.1 Design of gRNAs for CG6279 editing

Initial guidance for CRISPR editing in *Drosophila* was provided by Caitlin Anderson and Dr Robert Johnston (Dept. of Biology, JHU). The CG6279 gene (chromosome 3L, bases 11093371 to 11096245)[65] was analyzed for potential CRISPR sites using the 'CRISPR Optimal Target Finder' algorithm.[112] A pair of CRISPR target sites was selected for gene deletion (Appendix Fig. D1). A CRISPR target site closest to codon GAG encoding active site Glu-154 (Appendix Fig. D1) was selected for introduction of a single nucleotide mutation resulting in codon change to CAG and an amino acid substitution to Gln via HDR. The target nucleotide G (Appendix Fig. D1) was 18 bp upstream of the Cas9 cut site. CRISPR sites were chosen based on the low number of potential off-target sites (no more than 6) and at least 3 mismatches between the selected gRNAs and off-target sites (one of which is 12 or fewer bp upstream of the PAM) (Appendix Fig. D2). Sense and antisense oligonucleotides encoding the gRNAs were purchased from IDT and cloned into the pCFD3-U6:3-gRNA vector using published protocols.[113,114]

A ssODN repair template with 60-70 bp homology arms flanking the Cas9 cut sites was designed for HDR following gene deletion (Appendix Fig. D3, ssODN-1). An EcoRI restriction site was included between the flanking regions of the repair template for potential screening of mutants. The ssODN repair template for introducing the E154Q mutation included the targeted mutation as well as another silent mutation destroying the PAM site to prevent repeated Cas9 cleavage after HDR has occurred (Appendix Fig. D3, ssODN-2). The ssODN templates for HDR were purchased from IDT.

The pair of gRNA encoding vectors (250 ng/µl each) for targeted gene deletion and the ssODN template (100 ng/µl) were co-injected in *Drosophila* embryos expressing Cas9 in the germline (BDSC # 51323) by Bestgene Inc (Chino Hills, CA). The single gRNA encoding vector (500 ng/µl) and the ssODN repair template (100 ng/µl) for introducing the E154Q mutation were similarly co-injected in *Drosophila* embryos by Bestgene Inc.

5.2.2 Genetic experiments

All flies were raised on standard yeast, molasses, cornmeal and agar medium at 25°C. Larvae from surviving injected embryos were obtained from Bestgene Inc. Each injected embryo that developed into an adult was mated to *MKRS/TM6B* (balancers for 3rd chromosome[81]) *Drosophila* provided by the Xin Lab (JHU, Biology Dept.). 5 to 10 1st filial generation (F1) male progeny with the humeral phenotype (from the *TM6B* balancer) from each fertile cross were individually mated again with *MKRS/TM6B* balancer females. Male and female progeny from the 2nd filial generation (F2) with the humeral phenotype were mated to each other to establish individual *Drosophila* stocks. The 3rd chromosome (non-balancer) in a particular stock has been inherited from the same F1 parent. Single homozygous (*non-TM6B*) 3rd filial generation (F3) males were screened by PCR for targeted mutations.

5.2.3 Screening for mutations induced by CRISPR

A single male from each stock was evaluated for desired the mutation by PCR. The male was homogenized in 50 µl of buffer A (25 mM NaCl, 1mM EDTA, 10 mM TRIS pH 8.2). Proteinase K (1 µl of 20 mg/ml stock, Thermo Scientific) was added to the homogenate and incubated for 1 hr at 37 °C after which homogenate was incubated at 98-100 °C for 1 to 2 mins. An aliquot of homogenate (4 µl) was used as a template for a PCR

reaction using *Taq* DNA polymerase (Thermo Scientific) wherein forward primer 5'–ACACTTTGCCGATCTTGGCG-3' and reverse primer 5'-CAGTTCAAAGGAGTG-GTGACC-3' were used to amplify a 561 bp region spanning the dehalogenase domain of CG6279. Potential gene deletion mutants ($CG6279^{Del}$) were further analyzed by flanking PCR with forward primer 5'–TCTCGCTCGCTCTTTTACTGCA-3' and reverse primer 5'-CTTACCGTAGCTATCGTTATCGA-3' followed by sequencing of the PCR product (Genewiz).

For the E154Q mutant ($CG6279^{E154Q}$), successful mutation of the codon from GAG to CAG introduces a PvuII restriction site (CAG^CTG). The 561 bp PCR product described above yields smaller fragments (301 bp and 260 bp) upon restriction digestion with PvuII upon if the codon is mutated. Identification of the $CG6279^{E154Q}$ mutant was subsequently performed by restriction digestion of the 561 bp PCR fragment described above with PvuII (NEB) prior to agarose gel electrophoresis. Appearance of fragments smaller than 561 bp (301 bp and 260 bp) indicated target mutation. Target mutation were confirmed by sequencing the undigested PCR products from potential mutants (Genewiz).

5.2.4 Male fertility testing

The balanced deficiency (*Df*) fly stock for CG6279 carrying a deletion on chromosome 3L (11070526 to 11247145) was obtained from the Bloomington *Drosophila* stock center (BDSC # 27578).[115] Homozygous virgin females of genotype $CG6279^{Del}$, $CG6279^{E154Q}$ and $CG6279^{+}$ (*yw* flies) were mated with *Df/TM6* males. Progeny males of genotype $CG6279^{Del}/Df$, $CG6279^{E154Q}/Df$ and $CG6279^{+}/Df$ (control) were tested for fertility and fecundity (number of offspring produced) as previously described.[116]

5.3 Results

5.3.1 Deletion of CG6279 by CRISPR/Cas9

Two distinct CRISPR genome editing strategies were adopted to generate the desired dehalogenase null mutant *Drosophila*. The first approach involved deletion of CG6279 from its genomic locus by employing 2 gRNAs, one targeting the 5'end and another targeting the 3' end of the gene (Appendix Fig D1). Targeted deletion of CG6279 was designed to eliminate most of its open reading frame including start codon. Individual fruitfly stocks were produced such that flies in each stock inherited the 3rd chromosome from a single F1 progeny of a founder (individual fertile adult arising from each injected embryo). Single homozygous flies from each of the 35 individual stocks (representing 14 founders) were screened by PCR. The absence of a 561 bp DNA fragment covering the dehalogenase domain of CG6279 indicated a potential gene deletion mutant. Three potential mutants were identified amongst the 35 stocks that were screened (Fig. 5-3A).

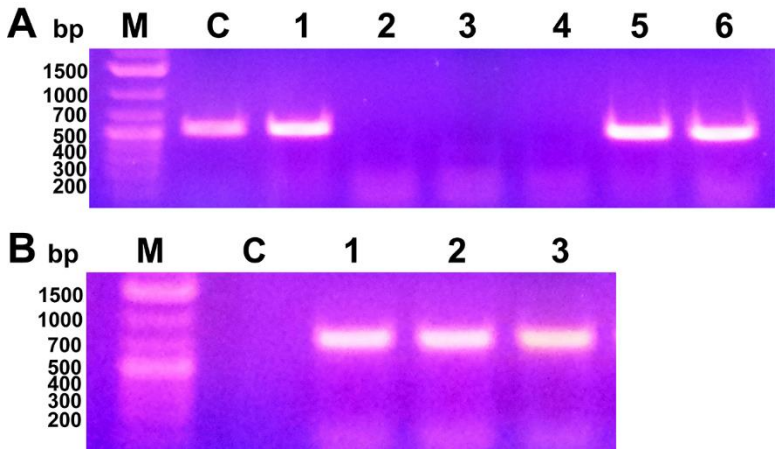

Figure 5-3. Agarose gel electrophoresis images for PCR products from screening for *CG6279^{Del}* mutants. (A) Screening of stocks for CG6279^{Del} mutants. Lanes 2, 3 and 4 are potential mutants (B) PCR products from potential mutants with primers flanking the CG6279 gene 467 bp upstream of start and 263 bp downstream of stop codon. M represents marker and C represents control (*Vasa-Cas9* strain).

Additional flanking PCR was performed for the potential mutants from the initial screen and the expected PCR product (777 bp) was observed (Fig. 5-3B). Gene deletion was verified by sequencing the flanking PCR product from mutant stock represented in lane 1 and this fly stock was used for further studies.

5.3.2 Mutation of active site Glu codon by CRISPR/Cas9

The second strategy involved introducing an amino acid substitution (E154Q) in the active site of the dehalogenase domain of CG6279 through a single mucleotide mutation at the Glu-154 codon (GAG to CAG, see in vitro characterization in chapter 3) that would greatly reduce the enzyme's catalytic efficiency. For screening, the same 561 bp PCR fragment described above was amplified as it covers the targeted codon. PCR products were digested with PvuII restriction endonuclease prior to agarose gel electrophoresis to identify mutants.

Screening of just 7 stocks derived from F1 progeny of distinct founders resulted in the identification of 3 potential mutants (Fig 5-4). PCR products from these potential mutants were sequenced to confirm the G to C mutation. Two of three mutants (lanes 4 and 5) showed an additional single nucleotide deletion 11 and 23 bases upstream of the targeted mutation respectively. These deletions were likely due to imprecise HDR and resulted in a codon reading frame shift at the point of deletion. The mutant indicated in lane 6 did not show any additional nucleotide insertions or deletions up to 278 bases upstream and 241 bases downstream of the targeted mutation. The fly stock represented by lane 6 was used for subsequent studies.

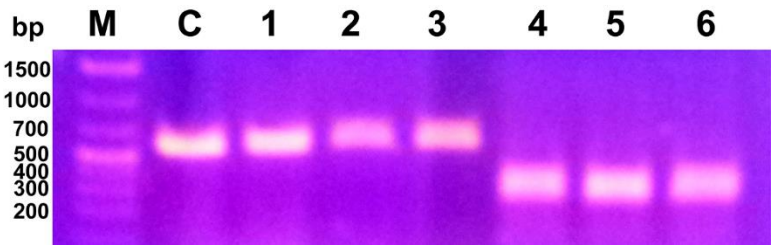

Figure 5-4. Agarose gel electrophoresis image for PCR products from screening for *CG6279*^{E154Q} mutants. Lanes 4, 5 and 6 are potential hits. M represents marker and C represents control (*Vas-Cas9* strain).

5.3.3 Fertility of CG6279 null mutant males

CG6279 null mutant males over a deficiency (*Df*) background were tested for fertility. *Df* strains contain defined deletions across a few or several genes on a particular chromosome.[117] *Drosophila* stocks containing these deletions are maintained as heterozygous and balanced stocks. *Df* strain 27578 from BDSC contains a deletion encompassing the CG6279 locus. A *Df* background facilitates assessment of individual alleles of a gene for phenotypic manifestations. It also enables the exclusion of any off-target chromosomal aberrations introduced through CRISPR/Cas9. *Drosophila* males of genotype *CG6279*^{Del}*/Df,* CG6279^{E154Q}*/Df* and *CG6279*$^{+}$*/Df* contain a single copy of either the targeted CG6279 deletion, the E154Q mutation at CG6279 or the wild-type CG6279 respectively. These males were individually mated to 3 *yw* females for a period of 5 days. The parent flies were then removed from the vials and the number of adult progeny produced (fecundity) was evaluated for each male fly. Of the 22 *CG6279*^{Del}*/Df* males tested, 20 were fertile while 20 of 21 *CG6279*^{E154Q}*/Df* males were fertile. All 16 *CG6279*$^{+}$*/Df* control males were fertile. The average number of progeny produced by *CG6279*^{Del}*/Df* males (102) and *CG6279*^{E154Q}*/Df* males (105) were similar to *CG6279*^{WT}*/Df* control males (100) (Fig. 5-5).

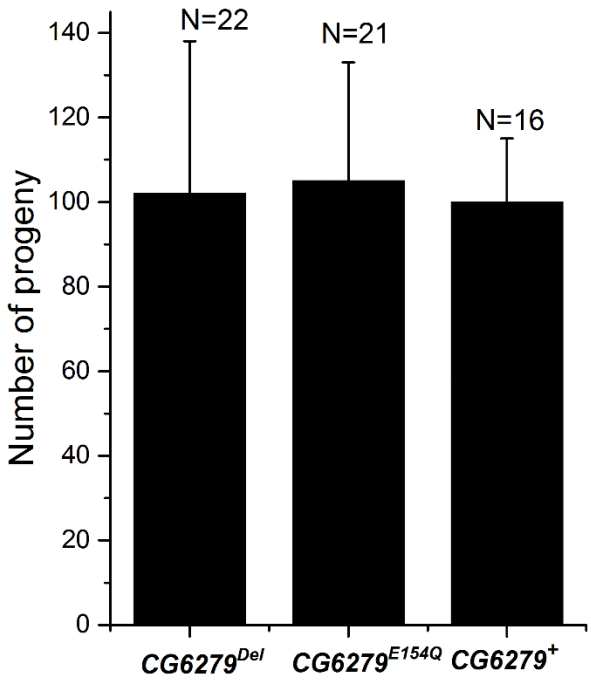

Figure 5-5. Average number of progeny produced by *CG6279^{Del}/Df*, *CG6279^{E154Q}/Df* and *CG6279^{+}/Df* (positive control) males when mated to 3 *yw* females over 5 days. N indicates total number of individual fertility testes and error bars represent standard deviation.

5.4 Discussion

CRISPR/Cas9 editing of the dehalogenase gene was efficient and the mutant strains were obtained through relatively few rounds of screening. All 3 *Drosophila* lines showing deletion of CG6279 derived from the same founder. This indicates that the CRISPR/Cas9 induced modification was successfully transmitted though the germline and stably inherited by the progeny. One of the 14 founders screened (7%) produced progeny with the targeted CG6279 gene deletion while one of 7 founders (14%) screened produced progeny with the targeted GAG to CAG point mutation within CG6279 without any additional aberrations. The aforementioned numbers do not reflect the actual efficiencies of CRISPR editing since screening of fly strains derived from additional founders was discontinued when the desired mutants were identified.

The dehalogenase gene CG6279 from *Drosophila melanogaster* encodes an additional N-terminal domain of unknown function in addition to the dehalogenase domain and splice variants of the gene transcript are predicted to result in 2 protein isoforms (isoform A and isoform B) with and without the N-terminal domain (Appendix Fig. B1). This additional N-terminal domain could have a function distinct from that of the dehalogenase domain. The $CG6729^{Del}$ mutant strain lacks both domains whereas the $CG6279^{E154Q}$ mutant strain encodes a catalytically deficient dehalogenase domain. Comparison of the fertility of males of each of these genotypes should enable identification of the relative importance of each individual domain towards spermatogenesis in *Drosophila*.

Deletion of CG6279 from the *Drosophila* genome did not significantly alter the fertility of males. This was surprising since in situ hybridization of *Drosophila* male testes (Chapter 4) clearly indicated high levels of the CG6279 transcript with peak expression in the post-meiotic stages. Consistent with the CG6279 gene deletion, the E154Q mutation in the active site of the dehalogenase domain also did not significantly affect the fertility of males. This indicates that the dehalogenase enzyme is not essential for fertility of male *Drosophila* and by extension not critical for spermatogenesis. It is possible that an alternate pathway compensates for the loss of the dehalogenase function in spermatogenesis. Further investigation would be required to identify possible rescue pathways for loss of dehalogenase activity.

5.5 Summary

The CG6279 gene encoding the dehalogenase enzyme in *Drosophila* was successfully edited using CRISPR/Cas9 technology. A mutant *Drosophila* strain was generated with a deletion at the CG6279 locus spanning almost the entire open reading frame (2211 bp of

2274 bp) including the start codon (*CG6279*^{Del}). A second mutant strain was also generated with the active site Glu-154 codon (GAG) changed to a Gln codon (CAG) via homology directed repair rendering the dehalogenase catalytically deficient without affecting the additional N-terminal domain as seen in isoform A of the gene transcript (*CG6279*^{E154Q}). *Drosophila* males of genotype *CG6279*^{Del}*/Df*, *CG6279*^{E154Q}*/Df* and *CG6279*$^{+}$*/Df* (control) were tested for fertility by mating to wild-type (*yw*) females. Despite the high expression of CG6279 in *Drosophila* testes during the post-meiosis stages of spermatogenesis as demonstrated in Chapter 4, both *CG6279*^{Del}*/Df* and *CG6279*^{E154Q}*/Df* males displayed fertility similar to *CG6279*$^{+}$*/Df* males (control). This indicated that deletion of CG6279 (including the additional N-terminal domain and dehalogenase domain) as well as selectively rendering the dehalogenase domain of CG6279 catalytically deficient does not significantly impact fertility of males. The dehalogenase encoding gene is likely not critical for spermatogenesis in *Drosophila*.

Chapter 6: Probing the metabolic activity of the *Drosophila* halotyrosine dehalogenase

6.1 Introduction

Little is known about halotyrosine function in arthropods. Halotyrosines are incorporated into scleroproteins[63,118] and may be produced as a result of peroxidase activity during insect cuticle sclerotization.[64] These halotyrosines could exert toxicity through disruption of cellular pathways. I-Tyr competes with the natural substrate Tyr and inhibits tyrosine hydroxylase with inhibition constants (K_i) in the sub-micromolar range.[119,120] Tyrosine hydroxylase catalyzes the conversion of Tyr to 3,4-dihydroxyphenylalanine (DOPA), a precursor to catecholamine neurotransmitters (Fig. 6-1).[121]

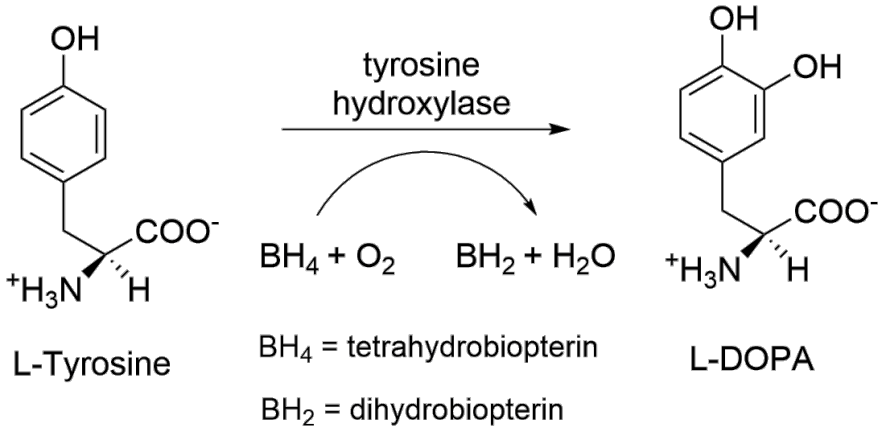

Figure 6-1. Reaction catalyzed by tyrosine hydroxylase.

Synthesis of the catecholamine neurotransmitter dopamine is well conserved between vertebrates and *Drosophila* and involves decarboxylation of DOPA by dopa decarboxylase to produce dopamine.[122] Dopamine has been demonstrated to modulate a number of behaviors in *Drosophila* including locomotion,[123] sleep and arousal,[124] memory and olfaction[125] and mating behavior.[126] Previously, I-Tyr has been used as an inhibitor to

deplete dopamine in neurological experiments involving *Drosophila* larvae.[127] Feeding of I-Tyr to *Drosophila* larvae at high concentrations (10 to 15 mg/ml of yeast paste media) was demonstrated to increase mortality, likely through inhibition of tyrosine hydroxylase since this effect could be rescued through co-feeding of L-DOPA.[128]

The *Drosophila* life cycle consists of four distinct stages.[129] The 1st stage is the egg or the embryonic stage. Within the fertilized egg, the embryo develops into the larva (2nd stage) and emerges from the egg in about 24 hrs post fertilization. Within the larval stage, there are 3 sub-stages or instars. The 1st instar (L1) larva hatches from the egg and this stage lasts for about 24 hrs.[130] The L1 larva then undergoes molting into the 2nd instar larva. The 2nd instar larva molts again into the 3rd instar larva. The larvae feed aggressively and undergo rapid growth and size increase from the 1st to the 3rd instar. The 3rd instar larva develops into the pupa (3rd stage). Within the pupa, the larval tissues undergo extensive remodeling. The adult (4th stage) finally emerges from the pupa.

The gene CG6279 encoding the halotyrosine dehalogenase is also expressed during the *Drosophila* pre-adult stages in the fat body (Chapter 4, Fig. 4-2), an organ compared to vertebrate liver for serving as the insect's metabolic center.[131,132] The dehalogenase may function in larvae to protect the organism from halotyrosine toxicity. In this chapter, we probe whether the dehalogenase gene provides a metabolic advantage to *Drosophila* through feeding halotyrosines to dehalogenase mutant strains ($CG6279^{Del}$ and $CG6279^{E154Q}$) and control strains ($CG6279^{WT}$ and *yw*) in the larval and adult stages of the life cycle.

6.2 Experimental Procedures

6.2.1 Materials and *Drosophila* media preparation

Giant apple juice (frozen) was purchased from Giant grocery. Methyl 4-hydroxybenzoate (methyl paraben or Tegosept) and erioglaucine disodium salt (FD&C blue 1) were purchased from Sigma Aldrich. Agar was purchased from Becton, Dickinson and Company. Stock solutions of 3-halotyrosines and Tyr were prepared in 0.1 N hydrochloric acid. Concentrations of 3-halotyrosine and Tyr stock solutions were determined based on extinction coefficients previously described.[63] Stock solution of Tegosept was prepared in ethanol. Stock solutions of sodium iodide (NaI), sodium bromide (NaBr) and FD&C blue 1 dye were prepared in water. Standard yeast, molasses, cornmeal and agar feeding medium (Appendix Table E1) was obtained from the JHU Biology Dept.

6.2.2 *Drosophila* strains

Homozygous dehalogenase mutant strains $CG6279^{Del}$ and $CG6279^{E154Q}$ generated through CRISPR editing as described in Chapter 5 were used in both larval and adult feeding studies. For generating the $CG6279^{WT}$ control strain, un-injected *vasa-Cas9* (BDSC # 51323)[133] flies were crossed with the *MKRS/TM6B* 3rd chromosome balancer strain similar to the procedure described in Chapter 5 (5.2.2). The *Drosophila* strain *yw*, commonly used as a reference strain in laboratory experiments[134,135] was employed as a control in addition to the $CG6279^{WT}$ strain in all feeding experiments and was obtained from Dr Xin Chen (JHU, Biology Dept.).

6.2.3 *Drosophila* larval halotyrosine feeding

For larval halotyrosine feeding experiments, mature egg laying adults of the desired genotype were transferred to vials containing apple juice agar medium (40% Giants apple

juice, 1.5% Agar and 0.05% Tegosept) and allowed to lay eggs. Adults were removed after 24 hrs and L1 larvae were collected between 24 to 30 hours after egg laying. Standard feeding media (Appendix Table E1) was supplemented with halotyrosine or control compounds and FD&C blue 1 dye (0.007% w/v). An equivalent amount of sodium hydroxide solution (previously titrated with 0.1 N hydrochloric acid) was added to feeding medium to neutralize the acid contributed by the 3-halotyrosine and Tyr stock solutions. A total of 50 L1 larvae were transferred to each individual vial. Food consumption by the larvae was detected through ingestion of the dye (Appendix Fig. E1). Vials were maintained at 25 °C and the number of larvae surviving to adulthood were quantified.

6.2.4 *Drosophila* adult halotyrosine feeding

One day old male and female adult flies (25 each) were transferred to a vial containing tissue (KimWipe, Kimberly Clark) soaked in 3 ml of apple juice-halotyrosine feeding mixture (75% v/v apple juice, 2 mM halotyrosine, 0.05% w/v methyl paraben and 0.05% w/v FD&C blue 1 dye and an equivalent amount of sodium hydroxide solution, previously titrated with 0.1 N hydrochloric acid, to neutralize the acid contributed by the 3-halotyrosine stock solutions). Consumption of liquid feeding mixture by adult flies was confirmed through ingestion of dye (Appendix Fig. E2). The adult flies were maintained on apple juice-halotyrosine mixture for at 25 °C for 5 days and number of surviving adults were quantified.

6.3 Results

6.3.1 *Drosophila* larval survival in halotyrosine containing feeding media

L1 larvae were raised on standard feeding medium supplemented with I-Tyr at 4 different concentrations of 100, 150, 200 and 300 µM respectively. A concentration dependent increase in I-Tyr toxicity was observed for larvae from all strains tested (Fig. 6-2). None of the larvae from the dehalogenase mutant strains ($CG6279^{Del}$ and $CG6279^{E154Q}$) survived to adulthood when raised on medium containing 300 µM I-Tyr. $CG6279^{WT}$ control larvae had a negligible average survival rate of 2% while larvae from the *yw* strain displayed the highest average survival rate of 18% at 300 µM I-Tyr. In medium containing 200 µM I-Tyr, $CG6279^{E154Q}$ larvae had a 0% survival rate which was significantly lower than average survival rates of both control strains, $CG6279^{WT}$ (28%, p-value <0.05) and *yw* (58%, p-value <0.01) (Fig. 6-2). However, $CG6279^{Del}$ larvae displayed a higher average survival rate (16%) than $CG6279^{E154Q}$ larvae at 200 µM I-Tyr. This was significantly lower than the average survival rate of *yw* larvae (p-value <0.01) but not significantly lower than $CG6279^{WT}$ larvae (p-value >0.05). At 150 µM I-Tyr, $CG6279^{Del}$ larvae displayed a high average survival rate (over 90%) that was very similar to the average survival rates of larvae from both the control strains. However, $CG6279^{E154Q}$ larvae displayed a significantly lower average survival rate (32%) compared to both $CG6279^{WT}$ and *yw* control larvae (p-value <0.01). Even at 100 µM I-Tyr, $CG6279^{E154Q}$ larvae displayed a lower average survival rate of 64% compared to larvae from the 3 other genotypes that showed average survival rates over 90%. Most of the larvae that progressed to pupation (≤95%) developed into adult flies indicating that increased mortality from I-Tyr toxicity was manifested prior to pupation.

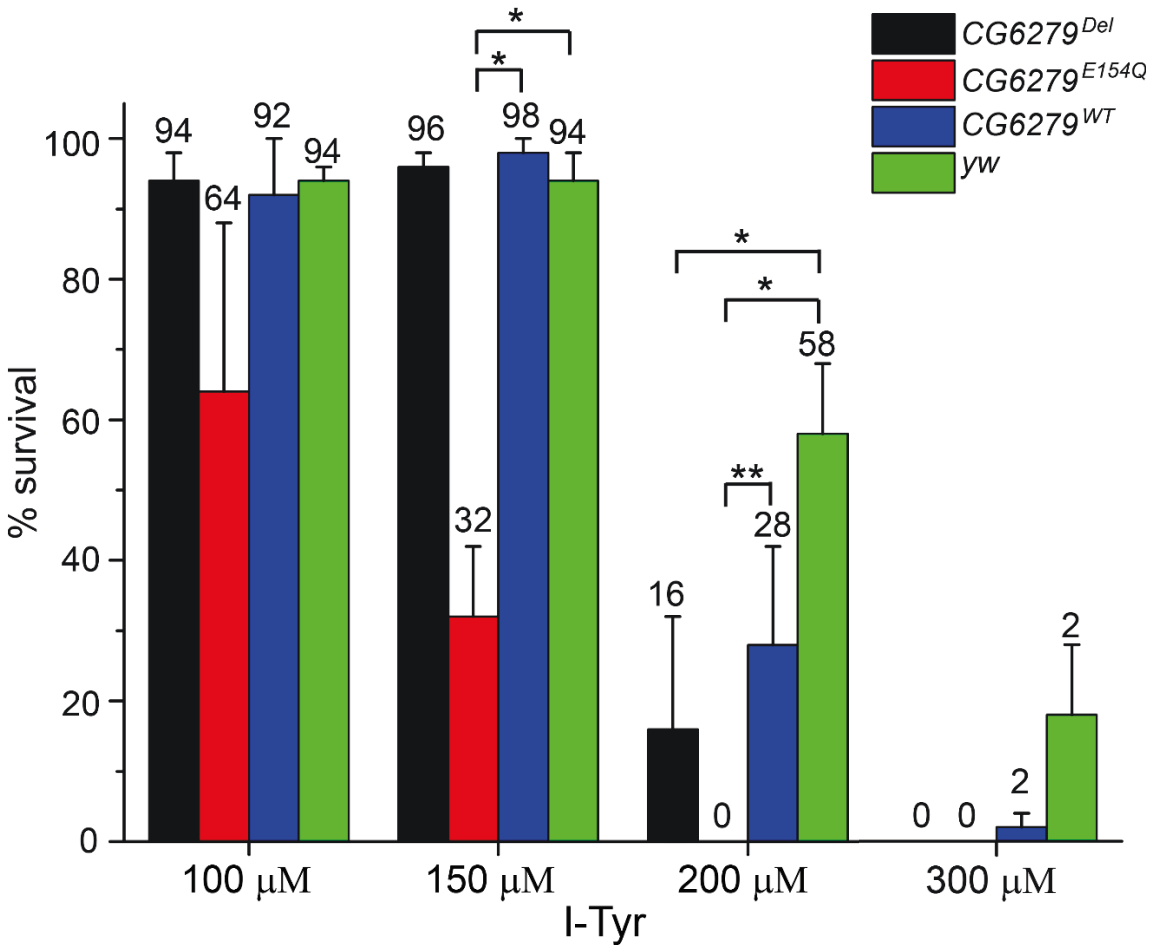

Figure 6-2. *Drosophila* larval survival rates when raised on media supplemented with I-Tyr. Indicated numbers represent averages from 3 or 4 independent measurements and error bars represent their standard deviations. (*) indicates p-value <0.01. (**) indicates p-value <0.05. Statistical significance was determined using Student's t-test for a two-tailed distribution with unequal variance. Bar plot is derived from data listed in Table E2.

Standard feeding medium was supplemented with Br-Tyr and Cl-Tyr at 1500 µM. Both these halotyrosines were tolerated at 5-fold higher concentration than the highest concentration of I-Tyr used indicating significantly lower toxicity than I-Tyr. Surprisingly, *CG6279*^{Del} larvae were most resistant to Br-Tyr toxicity with over 50% average survival to adulthood (Fig. 6-3A). Similar to I-Tyr, *CG6279*^{E154Q} larvae were most susceptible to Br-Tyr toxicity (8% survival). Both control strains (*CG6279*^{WT} and *yw*) containing the wild-

type dehalogenase gene displayed higher susceptibility to Br-Tyr compared to the *CG6279*^{Del} strain lacking the dehalogenase encoding gene.

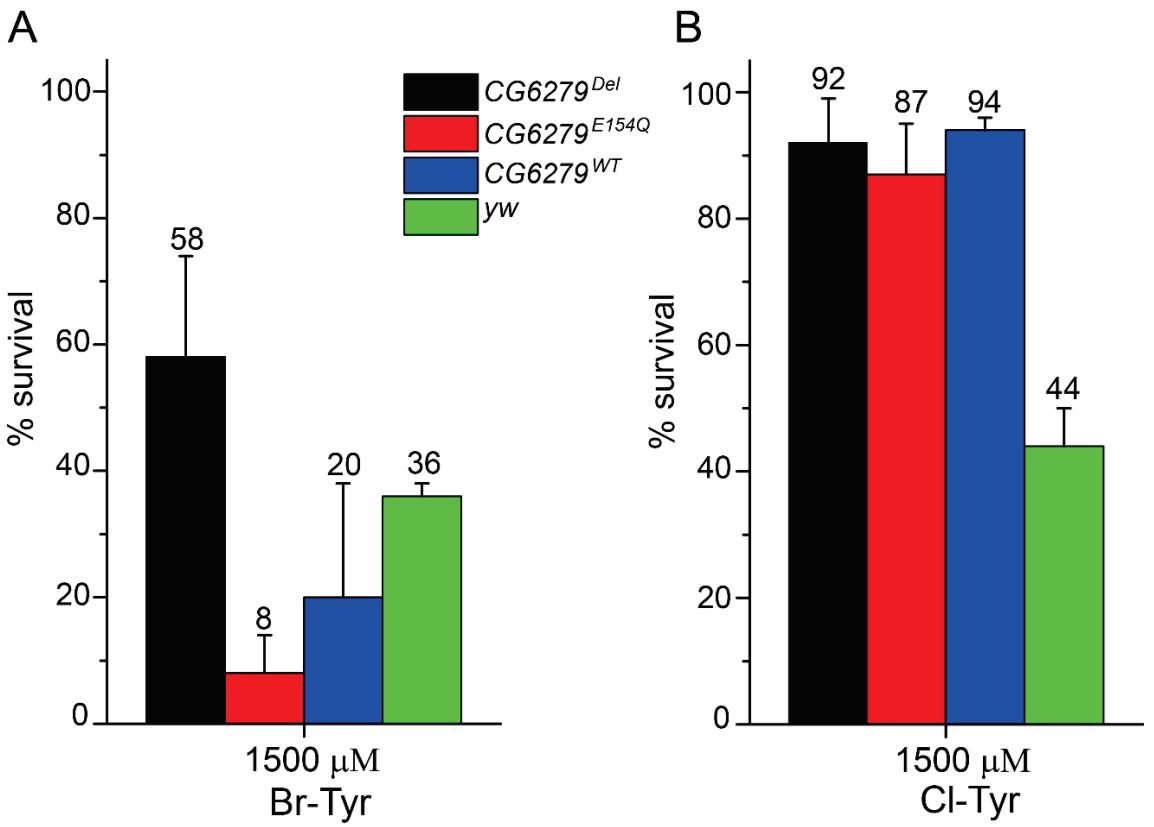

Figure 6-3. *Drosophila* **l**arval survival rates when raised on media supplemented with Br-Tyr (A) and Cl-Tyr (B). Indicated numbers represent averages from 3 independent measurements and error bars represent their standard deviations. Bar plot is derived from data listed in Table E2.

Larvae from *CG6279*^{Del}, *CG6279*^{E154Q} and *CG6279*^{WT} strains displayed high tolerance to Cl-Tyr with over 85% average survival for all three strains (Fig. 6-3B). However, the *yw* control strain displayed some susceptibility to Cl-Tyr toxicity with a 44% average survival rate. About 50% of the *yw* larvae that progressed to pupation did not produce adults. A similar effect was not observed for the other strains raised on Cl-Tyr.

Drosophila larvae from each of the aforementioned strains were also raised on medium containing the products of the dehalogenation reaction (Tyr, bromide and iodide) as controls. Tyr concentration was similar to the highest halotyrosine concentration added (1500 μM) while NaI and NaBr concentrations were similar to highest I-Tyr and Br-Tyr concentrations (300 μM and 1500 μM) respectively. Chloride was excluded due to its ubiquity in biological systems. The survival rates for larvae from the dehalogenase mutant strains as well as control strains exceeded 85% on medium supplemented with these control compounds indicating that these were not toxic to the larvae at concentrations comparable to the halotyrosines (Appendix Fig E3).

6.3.2 *Drosophila* adult survival in halotyrosine containing feeding media

One day old adult *Drosophila* (25 males and 25 females) were maintained for 5 days on apple juice containing halotyrosines at a high concentration (2 mM). A liquid feeding medium was used for the adults to facilitate higher intake through the proboscis compared to standard feeding media of soft solid consistency. $CG6279^{Del}$ and $CG6279^{E154Q}$ adult *Drosophila* exhibited high tolerance to all halotyrosines with over 80% average survival rates over 5 days. (Fig. 6-4). The survival of $CG6279^{Del}$ and $CG6279^{E154Q}$ mutant strains was comparable to both $CG6279^{WT}$ and *yw* control strains.

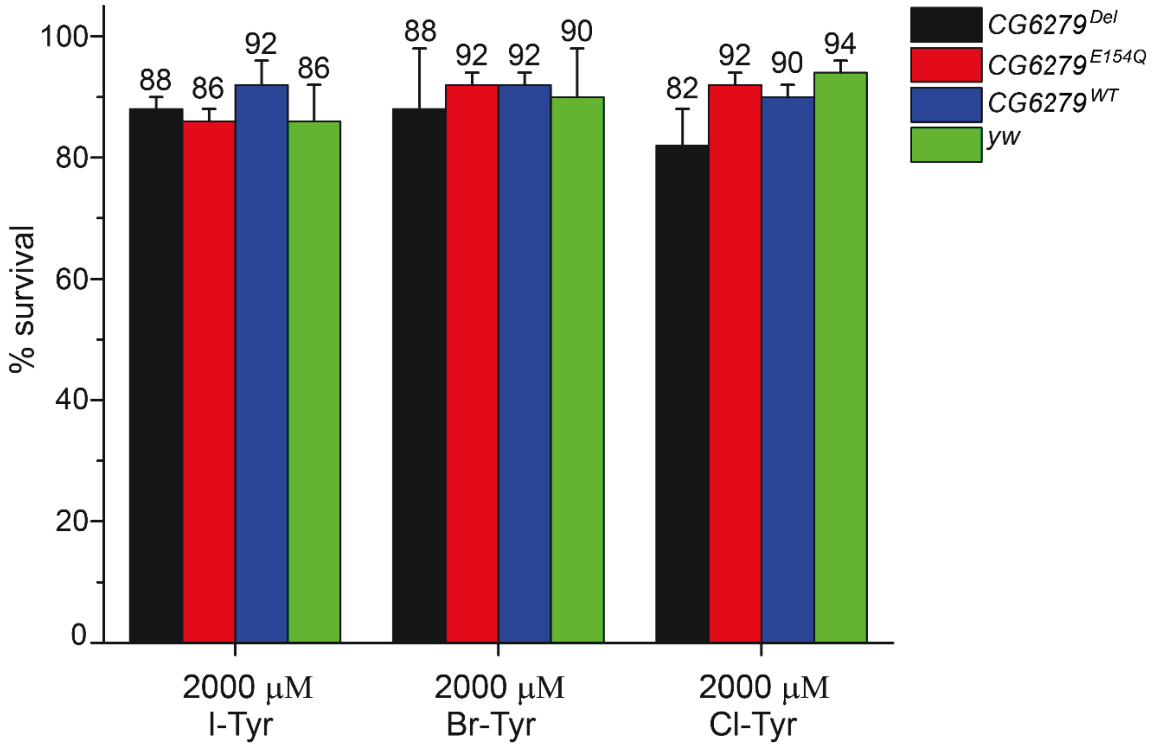

Figure 6-4. *Drosophila* adult survival rates when raised on apple juice media supplemented with halotyrosines. Indicated numbers represent averages from three independent measurements and error bars represent their standard deviations. Bar plot is derived from data listed in Table E3.

6.4 Discussion

The expression of the dehalogenase gene CG6279 in the pre-adult fat body suggested a possible metabolic role for the enzyme. Previous studies had indicated that I-Tyr inclusion in the feeding medium was toxic to *Drosophila* larvae through inhibition of tyrosine hydroxylase resulting in depletion of dopamine levels.[128] We rationalized that if the halotyrosine dehalogenase enzyme was expressed in a metabolic capacity, then both the $CG6279^{Del}$ and $CG6279^{E154Q}$ mutant larvae would be significantly more sensitive to halotyrosine toxicity relative to larvae from strains encoding the wild-type gene ($CG6279^{WT}$ and *yw*).

I-Tyr is the most lethal of the 3 halotyrosines when fed to *Drosophila* larvae while Cl-Tyr is the least toxic. Higher toxicity of I-Tyr may be due to stronger inhibition of tyrosine hydroxylase by I-Tyr compared to Br-Tyr and Cl-Tyr. However, toxicity through other mechanisms cannot be ruled out. Introduction of an E154Q mutation that renders the dehalogenase catalytically deficient results in significantly increased susceptibility of the larvae to I-Tyr toxicity compared to larvae from control strains containing the wild-type gene. Surprisingly, larvae with a genomic deletion of the dehalogenase gene were only modestly more sensitive to I-Tyr compared to a wild-type control. Variability in I-Tyr toxicity was also evident between the two control strains. Differential toxicity of I-Tyr towards $CG6279^{E154Q}$ and $CG6279^{Del}$ mutant larvae was unexpected given that larvae from both strains lack the ability to efficiently metabolize I-Tyr. It is possible that a catalytically deficient dehalogenase increases the sensitivity to I-Tyr through an off pathway effect relative to complete absence of the enzyme.

Very high concentrations of Br-Tyr (5-fold higher than the highest I-Tyr concentration tested) were required to induce toxicity. The average survival rate of $CG6279^{Del}$ larvae in Br-Tyr supplemented media was higher than the survival rates of larvae from control strains suggesting Br-Tyr at this high concentration may exert toxicity faster than it can be metabolized. Cl-Tyr was generally well tolerated at concentrations equivalent to Br-Tyr by larvae from both mutant strains and the $CG6279^{WT}$ control strain precluding any comparative analysis. Only larvae from the *yw* control strain displayed sensitivity to Cl-Tyr which is likely a peculiarity associated with that particular *Drosophila* strain.

Adult *Drosophila* from both mutant strains as well as the control strains did not display significant variation in survival rates (>80%) when maintained on apple juice containing

high concentrations each of halotyrosine for 5 days. This lack of toxicity is not so surprising for Br-Tyr and Cl-Tyr which were found to be significantly less toxic than I-Tyr in the larval stages. However, I-Tyr was also well tolerated by adult *Drosophila* at high concentration (2000 μM). It is possible that the adults are more resistant to I-Tyr toxicity than the larvae if dopamine signaling is more critical for survival of larvae. Alternatively, the physiological levels of I-Tyr upon feeding to adults may be lower compared to larvae feeding on I-Tyr containing media due to the minimal feeding by the adults relative to more aggressive feeding by the larvae prior to metamorphosis.

6.5 Summary

I-Tyr had been previously shown to be toxic to *Drosophila* larvae though inhibition of dopamine biosynthesis.[128] To probe whether the *Drosophila* halotyrosine dehalogenase is metabolically active in vivo, we performed feeding assays for dehalogenase null mutant *Drosophila* strains in the larval as well as the adult stage using media supplemented with halotyrosines. *Drosophila* strains encoding the wild-type dehalogenase gene were used as controls. Amongst the halotyrosines, I-Tyr was most toxic to *Drosophila* larvae in a dose dependent manner while Br-Tyr and Cl-Tyr were significantly less toxic. The $CG6279^{E154Q}$ mutant larvae encoding a catalytically deficient dehalogenase enzyme were most susceptible to I-Tyr toxicity. Survival rates for $CG6279^{E154Q}$ larvae were significantly lower than for larvae from both control strains. Surprisingly, $CG6279^{Del}$ larvae lacking the dehalogenase gene did not display significantly higher susceptibility to I-Tyr toxicity compared to larvae from the $CG6279^{WT}$ control strain. *Drosophila* adults from both CG6279 mutant as well as control strains were more tolerant of halotyrosines including I-

Tyr at high concentrations compared to the larvae. Based on significantly different susceptibilities to halotyrosine toxicity demonstrated by larvae from both dehalogenase null mutant strains, it is uncertain whether the enzyme functions in vivo in a metabolic capacity.

Chapter 7: Conclusions

Structural similarities between IYD and bacterial nitro-reductases were intriguing since IYD is the only mammalian member of a primarily prokaryotic enzyme superfamily. In order to determine whether the occurrence of IYD correlated with the need for iodide salvage, we defined an amino acid signature for deiodinase activity based on the co-crystal structure of mammalian IYD with its substrate (mmIYD·I_2-Tyr).[28] Conservation of the substrate interacting active site residues Glu-153, Tyr-157 and Lys-178 and FMN interacting Thr-235 were deemed to be indicative of IYD activity. We identified this amino acid signature in homologous proteins from diverse organisms including non-chordate invertebrates and prokaryotes, none of which are associated with TH synthesis nor have a thyroid gland. The kinetic characterization of selected IYD homologs from organisms belonging to diverse phyla within the animal kingdom as well as prokaryotic representatives revealed that these enzymes acted as deiodinases. This confirmed that substrate interacting residues Glu, Lys and Tyr as well as FMN interacting Thr are robust indicators of deiodinase activity. Presence of deiodinases in non-chordate invertebrates and prokaryotes suggested that the origin of IYD pre-dates the evolution of TH signaling.

Non-chordate invertebrates do not require iodide recycling for TH synthesis and alternative halotyrosines could be their primary substrates. The soluble domain of the IYD homolog from the model organism *Drosophila melanogaster* (dmIYD) was characterized in vitro to probe the scope of its catalytic activity as a representative of non-chordate invertebrates. Through substrate affinity and steady state kinetics, we found that I-Tyr, Br-Tyr and Cl-Tyr are all possible substrates for dmIYD in vitro and either one or all of these

may be physiologically relevant. Thus dmIYD may serve generally as a halotyrosine dehalogenase.

The amino acid indicators of IYD activity may be further refined through probing the relative contributions of the active site residues. Previous studies with mammalian IYD provided limited information regarding the relative importance of Glu-153, Tyr-157 and Lys-178 residues interacting with the substrate zwitterion.[29] Using dmIYD and its active site mutants we showed that the active site Glu residue is most critical for catalytic activity followed by the Lys residue while the Tyr residue did not significantly impact catalytic activity. Given the conservation of these residues in IYDs from mammals to prokaryotes, these findings should be broadly applicable to IYD homologs and further facilitate the annotation of potential IYDs within the nitro-FMN reductase superfamily. For example, putative NADH oxidases from *Hyperthermus butylicus* and *Staphylothermus marinus* (Accession WP_011822328.1 and WP_011839241.1 respectively) are likely deiodinases based on the conservation of the crucial Glu and Lys residues while the most dispensable Tyr residue is not conserved.

The biological function of halotyrosine dehalogenase in non-chordate invertebrates is unknown. *Drosophila* provided a promising model for in vivo investigations into its biological function. A peculiar feature of the dehalogenase gene from *Drosophila*, annotated as CG6279, was the presence of a large N-terminal domain of unknown function in addition to the dehalogenase domain. Splice variants of this gene with and without this N-terminal domain were also predicted. Previous genome-wide mRNA profiling studies in *Drosophila* indicated the highest expression of dehalogenase gene in the testes.[87,88] Through in situ hybridization experiments using *Drosophila* testes, we confirmed high

expression of CG6279. Peak expression appears in the post-meiotic stages of spermatogenesis. We attempted CG6279 gene silencing using RNAi in the testes to probe the effects of gene suppression on male fertility but these experiments were inconclusive since only one transgene (*eya-Gal4*) used to drive RNAi expression (*UAS-RNAi*) resulted in an abnormal testes phenotype and a similar testes phenotype was observed control males encoding *eya-Gal4* but not encoding *UAS-RNAi*. This phenotype was not reproduced by any other driver of RNAi expression further suggesting that the observed phenotype was likely due to the *eya-Gal4* transgene rather than CG6279 gene silencing.

Using CRISPR/Cas9 gene editing technology as an alternative approach, we successfully generated two CG6279 null mutant *Drosophila* strains. The $CG6279^{Del}$ strain had a deletion in the genome spanning the CG6279 gene including the N-terminal and the dehalogenase domains. The $CG6279^{E154Q}$ strain incorporated a point mutation resulting in an amino acid substitution in the active (E154Q) of the dehalogenase domain that greatly diminished the catalytic efficiency of the enzyme in vitro. Despite the high expression of the CG6279 gene during spermatogenesis, we found that both loss of function mutations did not significantly affect the fertility of the *Drosophila* males indicating that CG6279 is not critical for spermatogenesis. Alternatively, a rescue mechanism might be triggered in the testes that compensates for loss of dehalogenase function but further experimentation would be required to investigate this possibility.

CG6279 gene expression was also reported in the fat body during larval and pupal stages of the *Drosophila* life-cycle.[87,88] The insect fat body is homologous to the vertebrate liver serving as a center for carbohydrate, lipid and amino acid metabolism.[131,132] Expression of the dehalogenase in the fat body suggested a metabolic role. Previous studies had indicated

that ingestion of I-Tyr was toxic to *Drosophila* larvae due to inhibition of tyrosine hydroxylase, a critical enzyme in dopamine biosynthesis.[128] We reasoned that larvae from the $CG6279^{Del}$ and $CG6279^{E154Q}$ mutant strains would be more susceptible to halotyrosine toxicity compared to strains encoding the wild-type gene. I-Tyr was found to be more toxic to *Drosophila* larvae resulting in increased mortality at significantly lower concentrations compared to Br-Tyr and Cl-Tyr. Indeed, $CG6279^{E154Q}$ larvae were more susceptible to I-Tyr toxicity and displayed significantly lower survival rates relative to control strains. Surprisingly, these results were not mimicked by the $CG6279^{Del}$ strain which did not display significantly lower survival rates relative to the $CG6279^{WT}$ control strain. Based on these results it is unclear whether the dehalogenase gene functions in a metabolic capacity and needs to be further investigated. This difference between the $CG6279^{Del}$ strain and $CG6279^{E154Q}$ strain may be related to the function of the additional N-terminal domain and requires further investigation. Additionally, the halotyrosines may exert other physiological effects that do not result in increased mortality of organism. Further experiments to measure the impact of halotyrosine consumption on *Drosophila* fertility, locomotion and lifespan may reveal additional effects of these compounds.

The physiological function of the halotyrosine dehalogenase in *Drosophila* remains enigmatic. Our in vivo investigations of dehalogenase gene function were guided by mRNA expression profiles and previous studies investigating the effects of halotyrosines on *Drosophila*. The mRNA expression may not correlate with protein expression for the dehalogenase due to post transcriptional regulation. Experiments involving detection of the dehalogenase in *Drosophila* tissues using specific antibodies or proteomic analysis by mass spectrometry could provide evidence of protein expression. It is also possible that loss of

function mutations of CG6279 result in more subtle phenotypes that may have escaped detection.

Appendices

Appendix A: Supporting information for Chapter 2

Table A1. Extinction coefficients for IYD homologs.

IYD homolog	ε_{280} (M^{-1} cm^{-1})
dpIYD[a]	33920
nvIYD	34380
hhIYD	31400
pfIYD	37930
stIYD	25440
vhIYD	28420
tnIYD	36900

[a]Determined for truncated dpIYD (-55 amino acids, N-terminus)

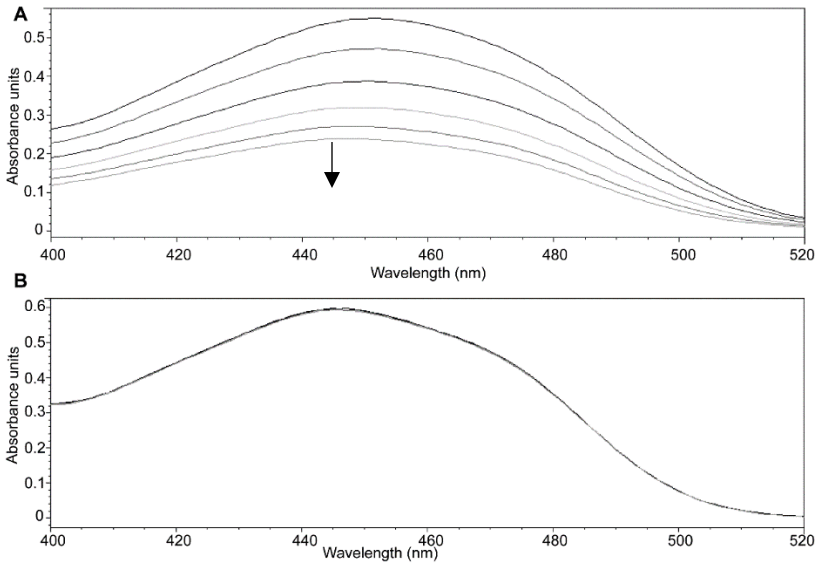

Figure A1. FMN depletion by BluB for detecting native activity was performed as previously described[47] at room temperature. (A) Represents reaction containing 50 µM FMN, 50 µM BluB, 300 µM NADPH 180 nM SSuE in 20 mM TRIS at pH 7.9. Depletion of FMN is reported at 445 nm every 150 secs (B) represents control without BluB.

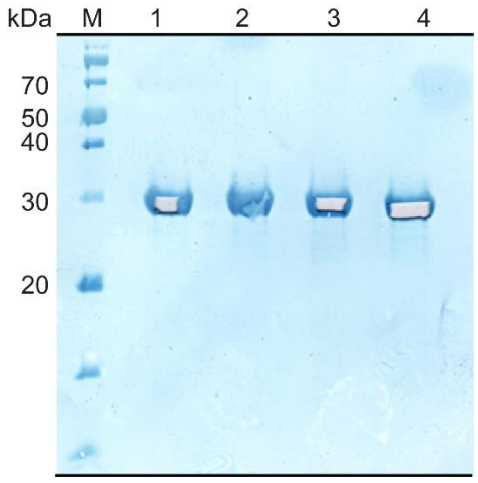

Figure A2. Coomassie stained polyvinylidene difluoride (PVDF) blot for purified dpIYD (lanes 1 to 4). Protein bands from lanes 3 and 4 were sent to JHMI core facility for N-terminal sequencing.

MENYVPFLVSHWRTVGLISVSIAAGVALGQLNQTHERAANKSQRVHTNSSIKAKQ⬇VKRE
SAIEQINEIEDDDEQSLEDQDVLPILFQYEKPSEADSIRRSEEFYRRMNQRRSVREISS
DPVALEVIENIIKTGGTSPSGAHTEPWTFVVVSNLEMKQQIRQIIEAEEEINYKQRMGD
VWVQDLQPVGTTWVKEYLTEAPWLILIFKQVHGFKRNGQKKIHYYNEISVSIATGFLLA
AIQEAGLVTVTTTPLNCGPSIRVLLGRPVNEKLLLLLPVGYPKVGATVPDFKRKPLHDI
MVHYQ

Figure A3. Identification of N-terminus of dpIYD. Sequencing results from JHMI indicated the N-terminus as MXXESAIEQINE. The first amino acid was reported as a methionine (M). The following 2 amino acids could not be determined but the next 9 amino acids matched the dpIYD sequence indicating the N-terminus likely begins at Valine 56 (arrow head). Amino acid sequence of the cleaved portion of the N-terminus indicated in red.

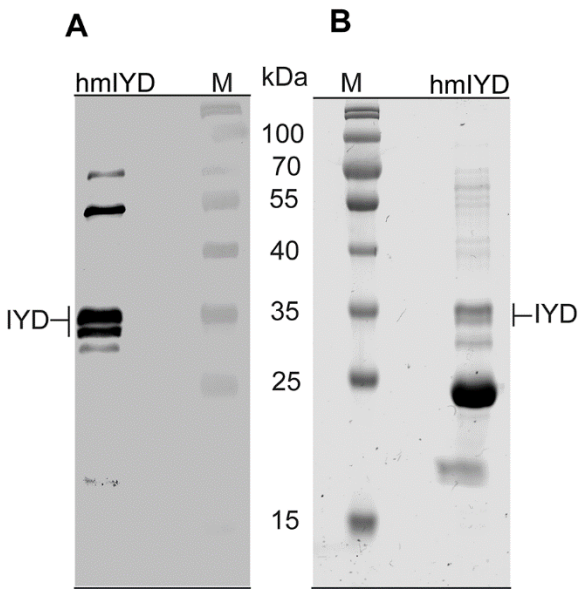

Figure A4. Identification of IYD from hydra (hmIYD). (A) Western blot of catalytically active protein fraction from hmIYD purification using His-tag mouse monoclonal antibody and goat anti-mouse antibody performed as per manufacturer procedures (Novagen). (B) Coomassie stained SDS-PAGE image of same active fraction tested in panel A.

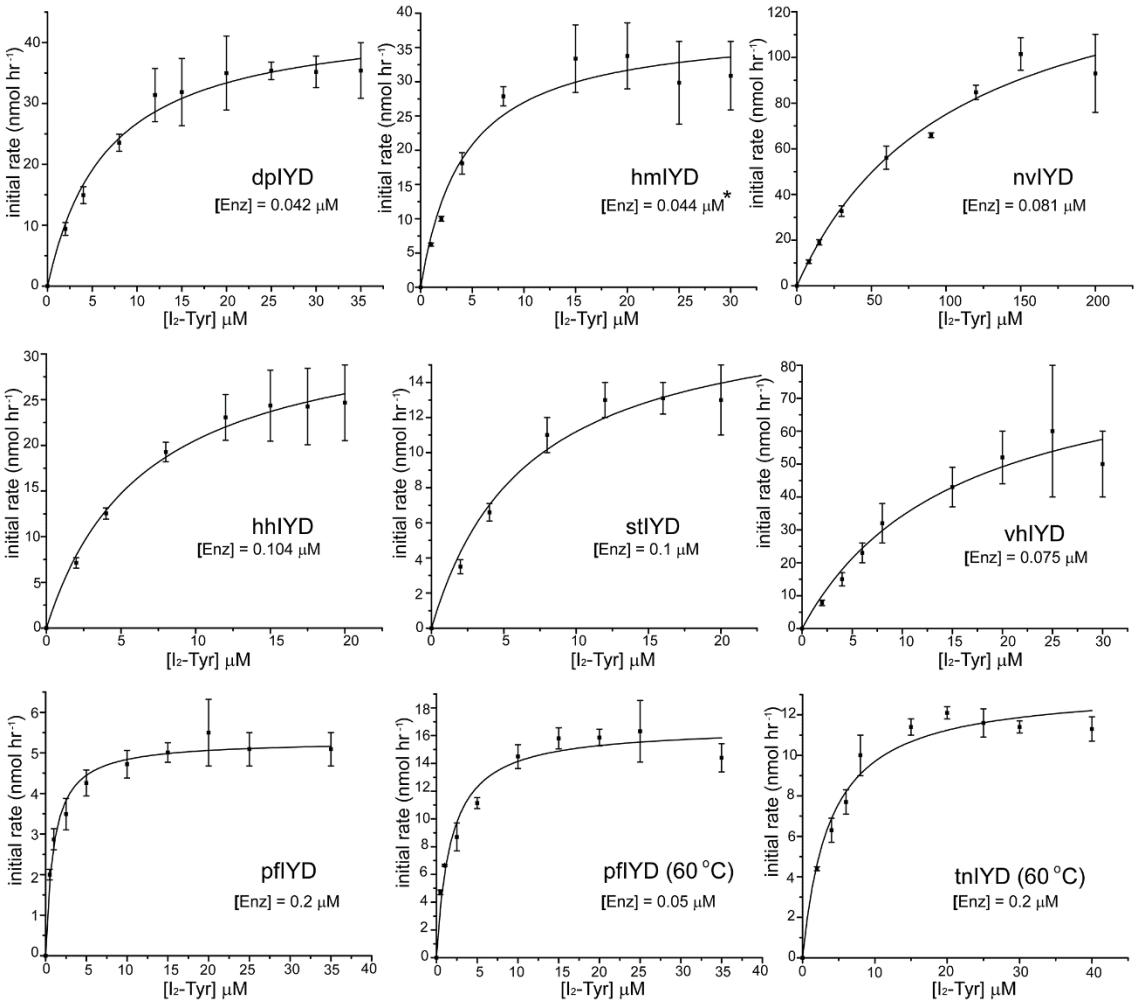

Figure A5. Steady-state analysis of ^{125}I-(I_2-Tyr) turnover by IYD homologs. Initial rates of deiodination were determined by quantifying $^{125}I^-$ released by enzyme. Each data point represents an average of 3 independent observations and the error bars represent the standard deviation from that average. Kinetic constants are derived from the best fit of the data (solid lines) to the Michealis-Menten equation. All experiments conducted at 25 °C unless mentioned otherwise. All nonlinear curve fitting was performed using Origin 6.0 or 7.0. Enzyme concentration represents the total protein concentration since addition of excess FMN is standard in these assays.[23,36]

Appendix B: Supporting information for Chapter 3

MEPKHRHCYFKRSASSCLVTPMHLHSSRISSSHSQSAGGIKPALASLDPPKTPNSIEKT
KKSGGFLSFLFGKRDRRARNQSETNGNHISLEDHDVRKLIRLYSKSVSELRSCLEELRT
LFEREYTIIERARRKCGEIMSPSIRYYNEFLFLVPHLNIHFDKDLPASGTSPLAAGKLL
QNPIANPEMLCQKLTNFTDFPLAGFPGNLLTKRPKKDPQDRPNKKIKPIPDQENAKQEN
PIEQEIQEAEKVADHEVSVEEKISGLEDQEQKDALSPSEQKNIQNSKDKTSHHHSNCKM
SSHEKDLVSTNLTCPCKPEVESSQSPVSCPWIAENSKCEPPCRETPPQSTRNSVSSPQN
SDISESSKEPAHLMEKEPSGQESANSQQVQMLCDMIRTELTTAPDFIFFDAKWRIIEIL
REVHKRQLVHQKAIPLNNTRRPIPPQKRKPGEPNQMLKNPKCMSDRSAHEKGQQRITMD
VDELISSSKLLKHWPSLFITLALIWIVKRLFFKGNRVLKTYNLDEQVEEEVEHFADLGD
ELQPALEDKPHVPFVPGQNLNPNGAKRLYELMRGRRSIRSFNSHPKPDLSVIEDCIRAA
GTAPSGAHTEPWTYCVVQEPELKRSIREIVEQEELVNYSQRMHPQWVTDLRPLQTNHVK
EYLTEAPYLILIFKQTYGLSENGKRMRRHYYNEISTSIAAGILLCALQAAGLASLVTTP
LNCGPALRNLLGRPVNEKLLILLPVGYPKDGCTVPDLARKNLSNIMVTF

Figure B1. Proteins resulting from IYD encoding gene CG6279. Isoform A (accession NP_648433.1) includes the N-terminal domain in red and the deiodinase domain in black. Isoform B (accession NP_001163414.1) resulting from alternative splicing of CG6279 includes just the deiodinase domain.

ATGAAAACTTATAATTTAGATGAACAAGTTGAAGAAGAAGTTGAACACTTTGCAGATTT
AGGAGATGAACTGCAGCCAGCCCTCGAAGATAAACCTCATGTTCCGTTCGTTCCCGGTC
AAAATTTAAACCCAAATGGTGCCAAACGCCTGTACGAACTTATGCGTGGCCGTCGCTCC
ATTCGTTCCTTTAATTCCCATCCCAAACCTGATCTTTCCGTCATTGAAGACTGTATTCG
CGCCGCCGGTACCGCCCCCTCCGGAGCCCACACCGAACCATGGACCTACTGTGTCGTAC
AAGAACCCGAACTCAAACGTAGCATTCGCGAAATTGTAGAACAAGAAGAATTAGTCAAT
TACTCCCAACGTATGCATCCACAATGGGTAACAGATTTACGCCCCCTCCAAACCAATCA
TGTAAAAGAATACTTAACCGAAGCCCCCTATTTAATTTTAATCTTTAAACAGACCTATG
GCCTGTCAGAAAACGGGAAACGTATGCGCCGTCATTACTACAACGAAATTTCCACATCC
ATTGCCGCTGGAATTTTACTCTGTGCCTTACAAGCTGCCGGACTCGCTTCCCTTGTTAC
AACCCCCTTAAACTGCGGCCCAGCCTTACGCAACCTTTTAGGACGCCCTGTTAACGAAA
AACTGCTTATTTTACTGCCCGTCGGATACCCCAAAGATGGTTGTACCGTACCAGATCTC
GCCCGCAAAAACCTCTCAAATATTATGGTAACCTTTCACCACCACCATCACCACTAA

Figure B2. *E. coli* expression optimized gene sequence for dmIYD synthesized by Blue Heron Biotechnology with 6X His-tag at C terminus (codons in red) followed by the stop codon.

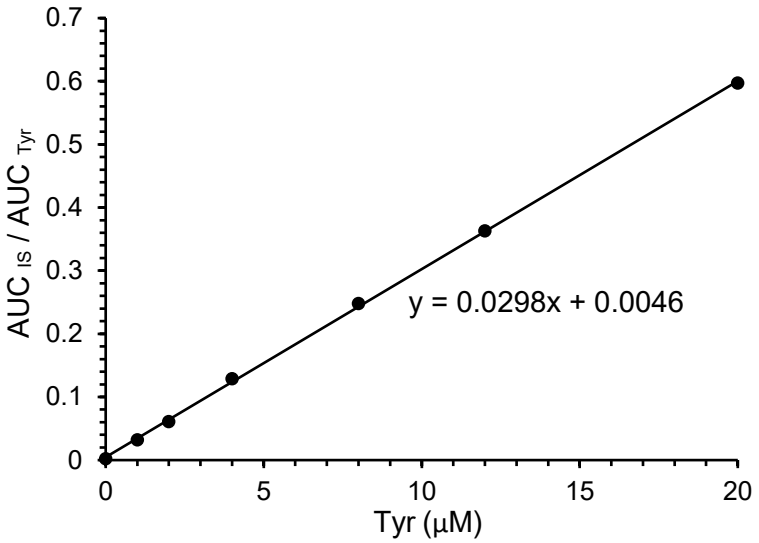

Figure B3. Standard curve for determination of Tyr produced by enzymatic turnover. Known concentrations of Tyr were spiked into enzymatic reaction mixtures lacking halotyrosine substrates and analyzed by HPLC. Ratio of area under curve (AUC) for m-cresol (IS) to AUC for Tyr was plotted against concentration of spiked Tyr. Each point represents the average of 2 individual measurements. Equation obtained from linear regression (indicated above) was used to calculate Tyr produced by enzymatic turnover of halotyrosines.

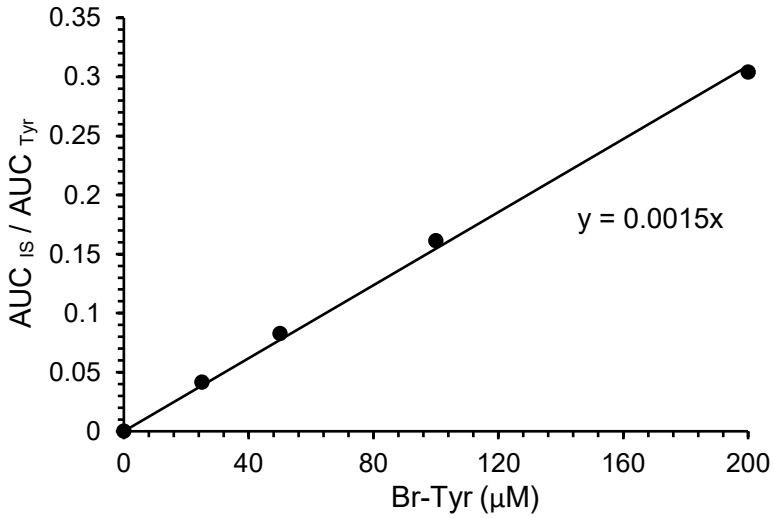

Figure B4. Standard curve for determination of Tyr impurity (not enzymatically generated) in bromotyrosine. Known concentrations of Br-Tyr were added to reaction mixtures lacking enzyme and analyzed by HPLC. Ratio of area under curve (AUC) for m-cresol (IS) to AUC for Tyr was plotted against concentration of Br-Tyr. Each point represents the average of 2 individual measurements. Equation obtained from linear regression (indicated above) was used to calculate Tyr impurity contribution to total Tyr produced.

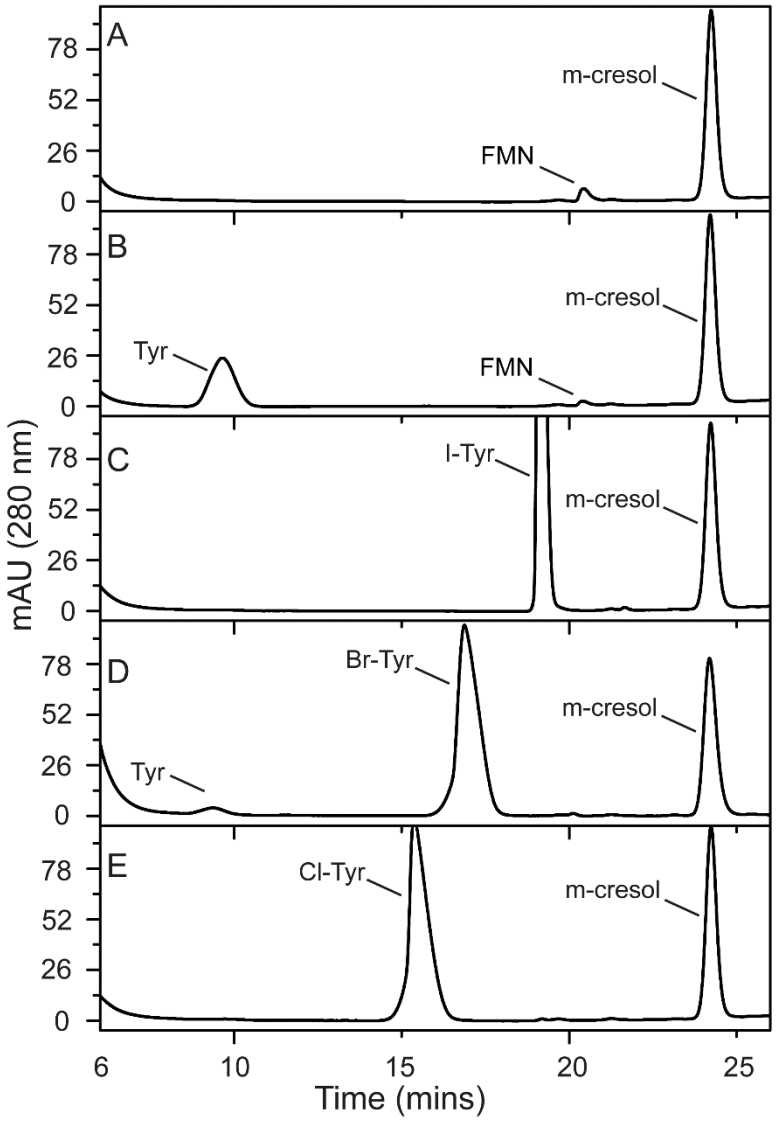

Figure B5. Standards for detection of Tyr after dmIYD turnover of the halotyrosines using reverse phase HPLC. All standards were incubated for 40 mins under conditions equivalent to the enzyme reactions i.e. 30 μM m-cresol, 200 mM KCl and 100 mM potassium phosphate pH 7.4 with addition of 0.5% dithionite in 5% sodium bicarbonate (100 μl) to a final volume of 1 ml followed by quenching with 88% formic acid (50 μl). (A) dmIYD only (0.2 μM). (B) Tyr (20 μM) and dmIYD (0.08 μM). (C) I-Tyr (40 μM). (D) Br-Tyr (40 μM). (E) Cl-Tyr (40 μM). FMN signal is derived from dmIYD.

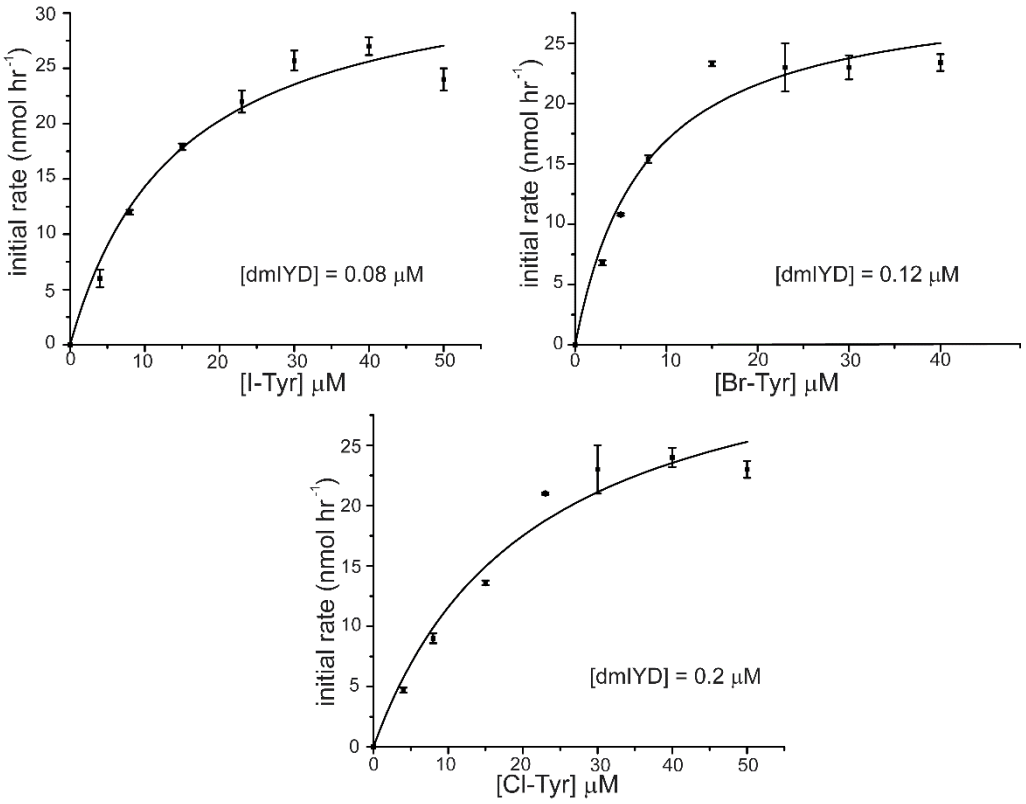

Figure B6. Steady-state analysis of halotyrosine turnover by dmIYD. Initial rate of dehalogenation was determined by quantifying Tyr product. Each data point represents an average of 3 individual observations and the error bars represent the standard deviation from those averages. Kinetic constants are derived from the best fit of the data (solid lines) to the Michealis-Menten equation. All nonlinear curve fitting was performed using Origin 6.0. [dmIYD] measured as associated FMN cofactor.

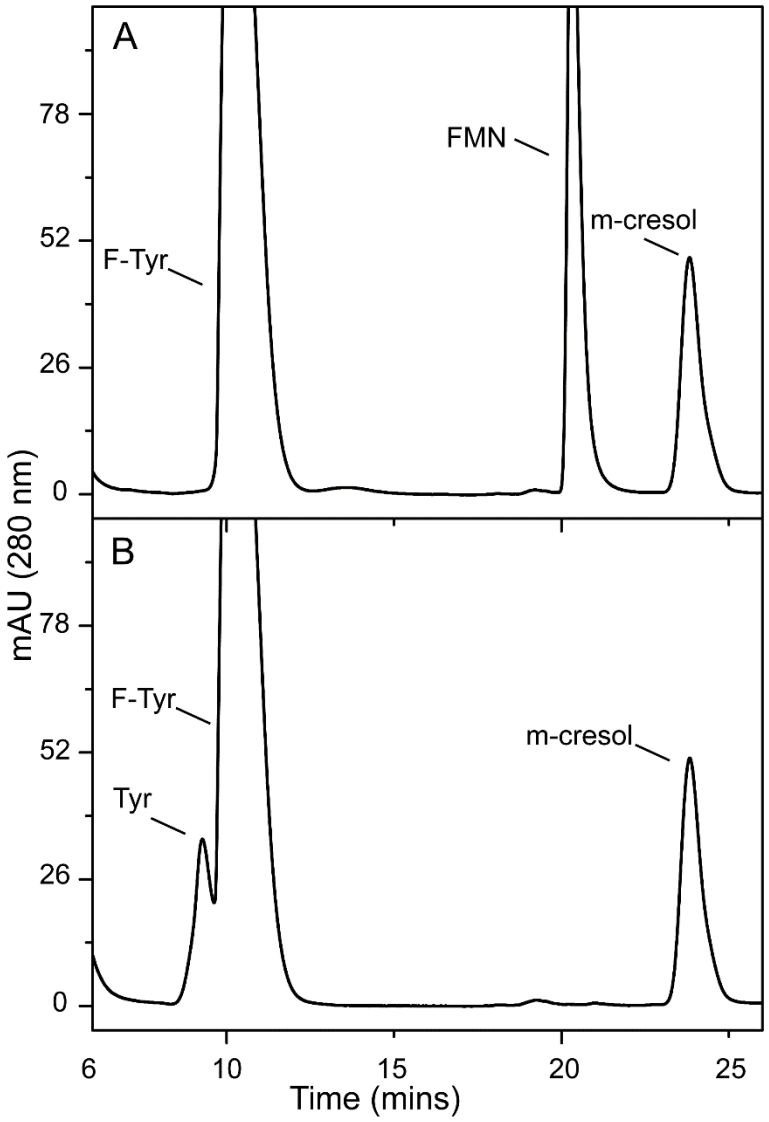

Figure B7. Potential defluorination of F-Tyr by dmIYD was monitored with reverse phase HPLC. Assay conditions were equivalent to those used for evaluating turnover of other halotyrosines i.e. 30 µM m-cresol, 200 mM KCl, 100 mM potassium phosphate pH 7.4 with addition of 0.5% dithionite in 5% sodium bicarbonate (100 µl) to initiate enzyme reduction followed by addition of 88% formic acid (50 µl) after 2 h to quench the reaction. (A) F-Tyr (200 µM) incubated with dmIYD (5 µM). (B) F-Tyr (200 µM) incubated with tyrosine standard (20 µM). The FMN signal is derived from dmIYD. No Tyr product was detectable from F-Tyr incubated with dmIYD.

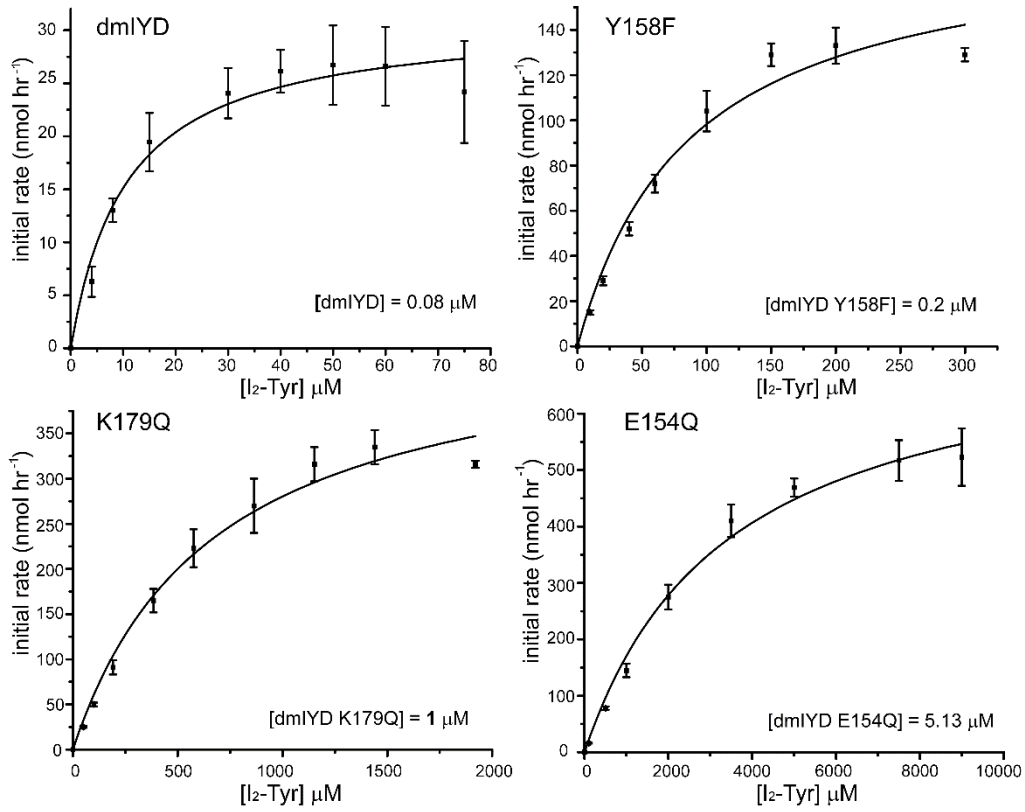

Figure B8. Steady-state analysis ^{125}I-(I_2-Tyr) turnover by dmIYD and its mutants. Initial rate of deiodination was determined by quantifying $^{125}I^-$ released by enzyme. Each data point represents an average of 3 independent observations and the error bars represent the standard deviation from that average. Kinetic constants are derived from the best fit of the data (solid lines) to the Michealis-Menten equation. All nonlinear curve fitting was performed using Origin 6.0. Enzyme concentration represents the total protein concentration since addition of excess FMN is standard in these assays.[23,36]

Appendix C: Supporting information for Chapter 4

AAAAAACTTCACAGTGACACTTACATTCACCCCAAAAAAAGTAAAAAAAAAATAGTCAGGATTACAAAAGT
TATATTTCCAGATAATGGTACAAAATTTTTTCGATTTATAATGCTGTTGCCATTAAAGAACCCAAAGACCG
TTGAAAGCGAGGTGAACCCGAATGTGACGCTACTGAAGCGCCAGATCGCCGAGGATGGAACCAAAACACAG
GCACTGTTACTTCAAGCGTTCCGCTTCGAGTTGCCTAGTGACGCCAATGCATCTCCACTCCTCCAGGATCA
GTTCCAGCCACAGTCAGTCAGCAGGAGGCATTAAGCCCGCTTTGGCATCACTAGATCCACCTAAAACACCC
AATAGCATCGAGAAAACAAAGAAATCCGGGGGATTTCTCTCTTTTTTGTTTGGCAAACGAGACAGGCGGGC
GAGGAATCAGAGTGAGACAAATGGAAACCATATCTCATTGGAGGATCATGATGTCAGGAAGCTTATACGTC
TCTATTGCAAGCATAATAGCCTTTACAACCGCCATAATCGCTACTATGGTAACAAAGATATCGACGATGAC
TGCTACAACAGAATGGTGCGCTCATTTCCGGGAAGAAGCGTCTCGGAACTGAGATCCTGTCTGGAGGAGCT
AAGGACGCTCTTTGAGCGAGAATACACCATTATCGAACGTGCTCGAAGGAAGTGTGGAGAGATAATGAGTC
CTTCAATTCGCTACTATAATGAATTTCTTTTCCTGGTGCCACATCTTAATATACACTTCGACAAGGATCTT
CCTGCCAGTGGAACTAGCCCACTCGCCGCCGGAAAACTTTTGCAAAATCCGATAGCCAATCCAGAGATGTT
GTGCCAAAAGTTGACCAATTTCACGGATTTTCCTCTGGCCGGTTTTCCCGGCAATCTACTAACCAAGAGGC
CCAAAAAAGACCCCCAAGACCGGCCAAACAAGAAGATCAAACCAATTCCTGATCAAGAAAATCCGAAGCAA
GAAAATCCCATAGAGCAGGAAATTCAGGAGGCAGAAAAAGTCGCTGATCATGAAGTTTCGGTAGAAGAGAA
AATCAGCGGGTTGGAGGATCAGGAACAGAAAGACGCACTCTCTCCATCGGAGCAGAAGAACATACAAAACA
GTAAGGATAAAACCAGTCATCATCATTCCAATTGCAAGATGTCCAGCCATGAGAAGGATTTAGTATCCACC
AATTTGACCTGTCCCTGCAAGCCGGAAGTGGAATCCTCACAAAGCCCAGTTTCCTGCCCCTGGATAGCCGA
AAATTCAAAGTGTGAACCCCCGTGTCGGGAAACACCACCACAATCTACGCGAAATTCCGTATCCAGTCCGC
AAAATTCAGATATTTCTGAATCCAGCAAAGAACCGGCTCACTTGATGGAGAAGGAACCATCCGGTCAGGAA
TCCGCGAATACCCAGCAAGTGCAGATGCTTTGCGATATGATACGTACCGAGCTAACCACTGCTCCCGATTT
TATATTCTTCGATGCCAAGTGGCGGATCATCGAGATTCTGAGAGAGGTGCACAAGCGCCAATTGGTTCACC
AGAAGGCCATTCCACTCAACAATACACGCCGACCGATTCCGCCGCAAAAACGAAAACCGGGTGAGCCCAAT
CAGATGCTTAAGAATCCAAAATGCATGTCGGATAGGAGTGCCCACGAAAAGGGACAACCGAGGATGTGCGA
GCATCACAAGGGCAGCCATTGCACCTACTGCTGCCGGAACAACAACTGAAGTGGAAACCAAAACCCATCCA
CCAGCATCCATCGCTATCAGCAATAGCCGTGACCCACTTAACCAAGTGATAAGCCAATCCGAGCGCAAAAG
CAAAGCCACTTTCTCTCCGTCGCAAAAGCAGTTACATCTTGAAGCACCATGGATGTGGACGAACTGATCAG
CAGCTCCAAGCTACTCAAGCATTGGCCAAGTTTATTTATCACTCTGGCGCTGATATGGATCGTAAAAAGAT
TGTTTTTTAAAGGGAATCGCGTATTAAAAACATATAATTTAGATGAACAAGTCGAGGAGGAGGGTAAGTTT
TAAGATTATCCACGGTGAACATAAAGCTATTCTATTCTATTTTCCAGTTGAACACTTTGCCGATCTTGGCG
ATGAACTGCAACCGGCTCTGGAGGATAAGCCCCATGTGCCCTTCGTTCCCGGCCAGAATCTCAACCCTAAT
GGAGCCAAACGCTTATATGAACTCATGCGTGGACGCCGTAGCATTCGCTCTTTTAGTAGCCATCCCAAACC
GGATCTCAGCGTGATAGAGGACTGCATTCGGGCGGCGGGAACAGCTCCCAGTGGCGCACACACAGAACCCT
GGACGTACTGCGTCGTACAGGAGCCGGAACTCAAGCGATCCATTCGGGAAATTGTGGAACAGGAGGAGCTG
GTCAACTACAGCCAGAGGATGCATCCGCAGTGGGTCACCGATCTGCGACCCCTGCAAACCAATCATGTCAA
GGAGTATCTCACCGAGGCACCATACCTCATACTGATCTTCAAGCAGACCTATGGATTGTCGGAGAATGGGA
AGCCAAGGAGGAGGCATTACTACAACGAGATATCCACTTCGATTGCGGCAGGTATCCTTTTGTGTGCCCTT
CAAGCGGCAGGACTCTCATCGCTGGTCACCACTCCTTTGAACTGTGGTCCAGCCTTGAGAAACCTGCTCGG
TCGACCCGTCAATGAAAAATTACTCATCCTACTGCCCGTTGGTTATCCGAAAGATGGATGCACAGTTCCCG
ACTTGGCGAGGAAGAATTTGTCAAATATTATGGTTACTTTTTAAGGTTTATTTACGTCCGTGCAGCTTACA
AAATAGTTCAATAAAACCAAACTGAAATTTATGGT

Figure C1. Sequence of CG6279 (genomic locus 3L, 11093371 to 11096245). The region targeted by *UAS-105378 RNAi* (isoform A) is indicated in yellow. The region targeted by *UAS-37267 RNAi* (isoform A and B) is indicated in blue. Both RNAi expressing fly strains were obtained from VDRC. Start and stop codons of gene are indicated in grey.

AAAAAACTTCACAGTGACACTTACATTCACCCCAAAAAAAGTAAAAAAAAAATAGTCAGGATTACAAAAGT
TATATTTCCAGATAATGGTACAAAATTTTTTCGATTTATAATGCTGTTGCCATTAAAGAACCCAAAGACCG
TTGAAAGCGAGGTGAACCCGAATGTGACGCTACTGAAGCGCCAGATCGCCGAGGATGGAACCAAAACACAG
GCACTGTTACTTCAAGCGTTCCGCTTCGAGTTGCCTAGTGACGCCAATGCATCTCCACTCCTCCAGGATCA
GTTCCAGCCACAGTCAGTCAGCAGGAGGCATTAAGCCCGCTTTGGCATCACTAGATCCACCTAAAACACCC
AATAGCATCGAGAAAACAAAGAAATCCGGGGGATTTCTCTCTTTTTTGTTTGGCAAACGAGACAGGCGGGC
GAGGAATCAGAGTGAGACAAATGGAAACCATATCTCATTGGAGGATCATGATGTCAGGAAGCTTATACGTC
TCTATTGCAAGCATAATAGCCTTTACAACCGCCATAATCGCTACTATGGTAACAAAGATATCGACGATGAC
TGCTACAACAGAATGGTGCGCTCATTTCCGGGAAGAAGCGTCTCGGAACTGAGATCCTGTCTGGAGGAGCT
AAGGACGCTCTTTGAGCGAGAATACACCATTATCGAACGTGCTCGAAGGAAGTGTGGAGAGATAATGAGTC
CTTCAATTCGCTACTATAATGAATTTCTTTTCCTGGTGCCACATCTTAATATACACTTCGACAAGGATCTT
CCTGCCAGTGGAACTAGCCCACTCGCCGCCGGAAAACTTTTGCAAAATCCGATAGCCAATCCAGAGATGTT
GTGCCAAAAGTTGACCAATTTCACGGATTTTCCTCTGGCCGGTTTTCCCGGCAATCTACTAACCAAGAGGC
CCAAAAAAGACCCCCAAGACCGGCCAAACAAGAAGATCAAACCAATTCCTGATCAAGAAAATCCGAAGCAA
GAAAATCCCATAGAGCAGGAAATTCAGGAGGCAGAAAAAGTCGCTGATCATGAAGTTTCGGTAGAAGAGAA
AATCAGCGGGTTGGAGGATCAGGAACAGAAAGACGCACTCTCTCCATCGGAGCAGAAGAACATACAAAACA
GTAAGGATAAAACCAGTCATCATCATTCCAATTGCAAGATGTCCAGCCATGAGAAGGATTTAGTATCCACC
AATTTGACCTGTCCCTGCAAGCCGGAAGTGGAATCCTCACAAAGCCCAGTTTCCTGCCCCTGGATAGCCGA
AAATTCAAAGTGTGAACCCCCGTGTCGGGAAACACCACCACAATCTACGCGAAATTCCGTATCCAGTCCGC
AAAATTCAGATATTTCTGAATCCAGCAAAGAACCGGCTCACTTGATGGAGAAGGAACCATCCGGTCAGGAA
TCCGCGAATACCCAGCAAGTGCAGATGCTTTGCGATATGATACGTACCGAGCTAACCACTGCTCCCGATTT
TATATTCTTCGATGCCAAGTGGCGGATCATCGAGATTCTGAGAGAGGTGCACAAGCGCCAATTGGTTCACC
AGAAGGCCATTCCACTCAACAATACACGCCGACCGATTCCGCCGCAAAAACGAAAACCGGGTGAGCCCAAT
CAGATGCTTAAGAATCCAAAATGCATGTCGGATAGGAGTGCCCACGAAAAGGGACAACCGAGGATGTGCGA
GCATCACAAGGGCAGCCATTGCACCTACTGCTGCCGGAACAACAACTGAAGTGGAAACCAAAACCCATCCA
CCAGCATCCATCGCTATCAGCAATAGCCGTGACCCACTTAACCAAGTGATAAGCCAATCCGAGCGCAAAAG
CAAAGCCACTTTCTCTCCGTCGCAAAAGCAGTTACATCTTGAAGCACCATGGATGTGGACGAACTGATCAG
CAGCTCCAAGCTACTCAAGCATTGGCCAAGTTTATTTATCACTCTGGCGCTGATATGGATCGTAAAAAGAT
TGTTTTTTAAAGGGAATCGCGTATTAAAAACATATAATTTAGATGAACAAGTCGAGGAGGAGGGTAAGTTT
TAAGATTATCCACGGTGAACATAAAGCTATTCTATTCTATTTTCCAGTTGAACACTTTGCCGATCTTGGCG
ATGAACTGCAACCGGCTCTGGAGGATAAGCCCCATGTGCCCTTCGTTCCCGGCCAGAATCTCAACCCTAAT
GGAGCCAAACGCTTATATGAACTCATGCGTGGACGCCGTAGCATTCGCTCTTTTAGTAGCCATCCCAAACC
GGATCTCAGCGTGATAGAGGACTGCATTCGGGCGGCGGGAACAGCTCCCAGTGGCGCACACACAGAACCCT
GGACGTACTGCGTCGTACAGGAGCCGGAACTCAAGCGATCCATTCGGGAAATTGTGGAACAGGAGGAGCTG
GTCAACTACAGCCAGAGGATGCATCCGCAGTGGGTCACCGATCTGCGACCCCTGCAAACCAATCATGTCAA
GGAGTATCTCACCGAGGCACCATACCTCATACTGATCTTCAAGCAGACCTATGGATTGTCGGAGAATGGGA
AGCCAAGGAGGAGGCATTACTACAACGAGATATCCACTTCGATTGCGGCAGGTATCCTTTTGTGTGCCCTT
CAAGCGGCAGGACTCTCATCGCTGGTCACCACTCCTTTGAACTGTGGTCCAGCCTTGAGAAACCTGCTCGG
TCGACCCGTCAATGAAAAATTACTCATCCTACTGCCCGTTGGTTATCCGAAAGATGGATGCACAGTTCCCG
ACTTGGCGAGGAAGAATTTGTCAAATATTATGGTTACTTTTTAAGGTTTATTTACGTCCGTGCAGCTTACA
AAATAGTTCAATAAAACCAAACTGAAATTTATGGT

Figure C2. Sequence of CG6279 (genomic locus 3L, 11093371 to 11096245). The region targeted by *UAS-shmiRNA-A* (isoform A) is indicated in yellow. The region targeted by *UAS-shmiRNA-AB* (isoforms A and B) is indicated in blue. The region targeted by *UAS-shmiRNA-B* (isoform B) is indicated in green. Start and stop codons of the gene are indicated in grey.

Appendix D: Supporting information for Chapter 5

AAAAAACTTCACAGTGACACTTACATTCACCCCAAAAAAAGTAAAAAAAAAATAGTCAGGATTACAAAAGT
TATATTTCCAGATAATGGTACAAAATTTTTTCGATTTATAATGCTGTTGCCATTAAAGAACCCAAAGACCG
TTGAAAGCGAGGTGAACCCGAATGTGACGCTACTGAAGCGCCAGATCGCCGAGGATGGAACCAAAACACAG
GCACTGTTACTTCAAGCGTTCCGCTTCGAGTTGCCTAGTGACGCCAATGCATCTCCACTCCTCCAGGATCA
GTTCCAGCCACAGTCAGTCAGCAGGAGGCATTAAGCCCGCTTTGGCATCACTAGATCCACCTAAAACACCC
AATAGCATCGAGAAAACAAAGAAATCCGGGGGATTTCTCTCTTTTTTGTTTGGCAAACGAGACAGGCGGGC
GAGGAATCAGAGTGAGACAAATGGAAACCATATCTCATTGGAGGATCATGATGTCAGGAAGCTTATACGTC
TCTATTGCAAGCATAATAGCCTTTACAACCGCCATAATCGCTACTATGGTAACAAAGATATCGACGATGAC
TGCTACAACAGAATGGTGCGCTCATTTCCGGGAAGAAGCGTCTCGGAACTGAGATCCTGTCTGGAGGAGCT
AAGGACGCTCTTTGAGCGAGAATACACCATTATCGAACGTGCTCGAAGGAAGTGTGGAGAGATAATGAGTC
CTTCAATTCGCTACTATAATGAATTTCTTTTCCTGGTGCCACATCTTAATATACACTTCGACAAGGATCTT
CCTGCCAGTGGAACTAGCCCACTCGCCGCCGGAAAACTTTTGCAAAATCCGATAGCCAATCCAGAGATGTT
GTGCCAAAAGTTGACCAATTTCACGGATTTTCCTCTGGCCGGTTTTCCCGGCAATCTACTAACCAAGAGGC
CCAAAAAAGACCCCCAAGACCGGCCAAACAAGAAGATCAAACCAATTCCTGATCAAGAAAATCCGAAGCAA
GAAAATCCCATAGAGCAGGAAATTCAGGAGGCAGAAAAAGTCGCTGATCATGAAGTTTCGGTAGAAGAGAA
AATCAGCGGGTTGGAGGATCAGGAACAGAAAGACGCACTCTCTCCATCGGAGCAGAAGAACATACAAAACA
GTAAGGATAAAACCAGTCATCATCATTCCAATTGCAAGATGTCCAGCCATGAGAAGGATTTAGTATCCACC
AATTTGACCTGTCCCTGCAAGCCGGAAGTGGAATCCTCACAAAGCCCAGTTTCCTGCCCCTGGATAGCCGA
AAATTCAAAGTGTGAACCCCCGTGTCGGGAAACACCACCACAATCTACGCGAAATTCCGTATCCAGTCCGC
AAAATTCAGATATTTCTGAATCCAGCAAAGAACCGGCTCACTTGATGGAGAAGGAACCATCCGGTCAGGAA
TCCGCGAATACCCAGCAAGTGCAGATGCTTTGCGATATGATACGTACCGAGCTAACCACTGCTCCCGATTT
TATATTCTTCGATGCCAAGTGGCGGATCATCGAGATTCTGAGAGAGGTGCACAAGCGCCAATTGGTTCACC
AGAAGGCCATTCCACTCAACAATACACGCCGACCGATTCCGCCGCAAAAACGAAAACCGGGTGAGCCCAAT
CAGATGCTTAAGAATCCAAAATGCATGTCGGATAGGAGTGCCCACGAAAAGGGACAACCGAGGATGTGCGA
GCATCACAAGGGCAGCCATTGCACCTACTGCTGCCGGAACAACAACTGAAGTGGAAACCAAAACCCATCCA
CCAGCATCCATCGCTATCAGCAATAGCCGTGACCCACTTAACCAAGTGATAAGCCAATCCGAGCGCAAAAG
CAAAGCCACTTTCTCTCCGTCGCAAAAGCAGTTACATCTTGAAGCACCATGGATGTGGACGAACTGATCAG
CAGCTCCAAGCTACTCAAGCATTGGCCAAGTTTATTTATCACTCTGGCGCTGATATGGATCGTAAAAAGAT
TGTTTTTTAAAGGGAATCGCGTATTAAAAACATATAATTTAGATGAACAAGTCGAGGAGGAGGGTAAGTTT
TAAGATTATCCACGGTGAACATAAAGCTATTCTATTCTATTTTCCAGTTGAACACTTTGCCGATCTTGGCG
ATGAACTGCAACCGGCTCTGGAGGATAAGCCCCATGTGCCCTTCGTTCCCGGCCAGAATCTCAACCCTAAT
GGAGCCAAACGCTTATATGAACTCATGCGTGGACGCCGTAGCATTCGCTCTTTTAGTAGCCATCCCAAACC
GGATCTCAGCGTGATAGAGGACTGCATTCGGGCGGCGGGAACAGCTCCCAGTGGCGCACACACAGAACCCT
GGACGTACTGCGTCGTACAGGAGCCGGAACTCAAGCGATCCATTCGGGAAATTGTGGAACAGGAGGAGCTG
GTCAACTACAGCCAGAGGATGCATCCGCAGTGGGTCACCGATCTGCGACCCCTGCAAACCAATCATGTCAA
GGAGTATCTCACCGAGGCACCATACCTCATACTGATCTTCAAGCAGACCTATGGATTGTCGGAGAATGGGA
AGCCAAGGAGGAGGCATTACTACAACGAGATATCCACTTCGATTGCGGCAGGTATCCTTTTGTGTGCCCTT
CAAGCGGCAGGACTCTCATCGCTGGTCACCACTCCTTTGAACTGTGGTCCAGCCTTGAGAAACCTGCTCGG
TCGACCCGTCAATGAAAAATTACTCATCCTACTGCCCGTTGGTTATCCGAAAGATGGATGCACAGTTCCCG
ACTTGGCGAGGAAGAATTTGTCAAATATTATGGTTACTTTTTAAGGTTTATTTACGTCCGTGCAGCTTACA
AAATAGTTCAATAAAACCAAACTGAAATTTATGGT

Figure D1. CRISPR sites selected within CG6279 gene (genomic locus 3L, 11093371 to 11096245). The two CRISPR sites chosen for gene deletion are highlighted in blue and green. The 1st nucleotide of the active site Glu codon is indicated in pink. CRISPR site closest to active site Glu codon is indicated in yellow. All PAMs are highlighted in red. The start codon and the stop codon are highlighted in grey. The gRNAs recognize the antisense strand for CRISPR sites indicated in blue and green and the sense strand for CRISPR site indicated in yellow. Cas9 cut sites are indicated by arrowheads.

CRISPR target	Chromosomal location	Target Direction	Status
GCTTCAGTAGCGTCACATTC GGG	3L:11096065..11096087	Anti-sense	On target
GCTGAAGCAGCGGCACATTC CGG	X:6826309..6826331	Sense	Off target
ATTTTAGTAGCGTTACATTC CGG	3L:23182180..23182202	Sense	Off target
GTTTTACTAGCGTCACTTTC GGG	2R:7104845..7104867	Anti-sense	Off target
GCAGCAGCAGCGTCACGTTC CGG	X:5683924..5683946	Anti-sense	Off target
GCTTGCTGGGCGTCACATTC TGG	X:6324699..6324721	Sense	Off target
GGAACTGTGCATCCATCTTT CGG	3L:11093479..11093501	Anti-sense	On target
ATAATTGTGCATCCATCTTG TGG	2R:8105299..8105321	Anti-sense	Off target
TCAGCTGACCATCCATCTTT AGG	2L:8518308..8518330	Sense	Off target
GCGAATGAGCTTCCATCTTT CGG	3R:31530841..31530863	Sense	Off target
TTAATTGGGCATCAATCTTT CGG	3L:16887531..16887553	Anti-sense	Off target
TGCAAGGTGCATCCATCGTT AGG	X:21420176..21420198	Anti-sense	Off target
GGAACTTTGCCTGCATCTTT GGG	2L:6454470..6454492	Anti-sense	Off target
AGCTGGTCAACTACAGCCAG AGG	3L:11093814..11093836	Sense	On target
AGGGAGTCATCTACAGCCAG GGG	3L:20417543..20417565	Anti-sense	Off target
CGCTGCACAACTACATCCAG AGG	3R:30853839..30853861	Anti-sense	Off target
TGCAGAACATCTACAGCCAG GGG	3R:30853839..30853861	Anti-sense	Off target

Figure D2. Analysis of potential off-targets for selected CRISPR sites.[112] All mis-matches with CRISPR gRNA are indicated in red. The PAM is separated from the CRISPR target by a space. The highlighted sequences correspond to those shown in Fig. A13

ssODNs	Sequence 5' to 3'
1	ATTTATAATGCTGTTGCCATTAAAGAACCCAAAGACCGTTGAAAGCGAGGTGAACCCGAAG AATTCGATGGATGCACAGTTCCCGACTTGGCGAGGAAGAATTTGTCAAATATTATGGTTAC TTTTTAAGGTTTAT
2	TACTGCGTCGTACAGGAGCCGGAACTCAAGCGATCCATTCGGGAAATTGTGGAACAGGAGC AGCTGGTCAACTACAGCCAGAGAATGCATCCGCAGTGGGTCACCGATCTGCGACCCCTGCA AACCAATCATGTCAAGGAGTAT

Figure D3. Single-stranded oligonucleotide templates (ssODN) for homology-directed repair (HDR) used in conjunction with gRNAs. Oligonucleotide 1 is designed for repair following gene deletion by cleavage at 5' and 3' CRISPR sites (blue and green sites indicated in Fig. D1). EcoRI restriction site for potential screening is indicated in red. Oligonucleotide 2 is designed to introduce a G to C (red) mutation changing a Glu codon to Gln. A silent mutation destroying PAM is indicated in grey. Additional highlighted sequences correspond to partial or complete CRISPR sites shown in Fig. D2.

Appendix E: Supporting information for Chapter 6

Table E1. Recipe for standard media used for larval feeding.

Ingredient	**Quantitiy**
Molasses	6.6% v/v
Yeast	2.7% w/v
Cornmeal	5.8% w/v
Agar	1.0% w/v
Propionic acid	0.5% v/v
Tegosept (10 % w/v in ethanol)	1.0% v/v
Water	To final volume

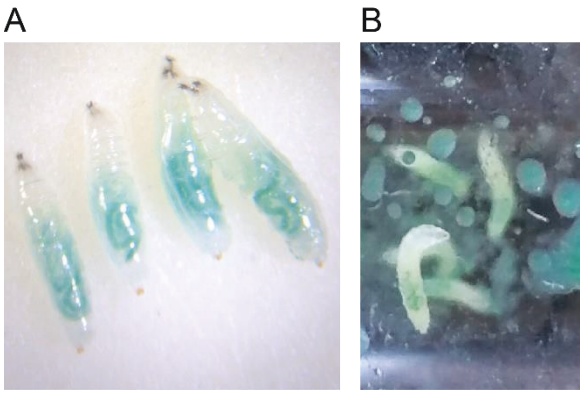

Figure E1. Representative image for *Drosophila* larvae raised on media containing halotyrosines and FD&C blue 1. (A) Third instar larvae that have consumed dye labeled media. (B) Third instar larvae on vial walls prior to pupation.

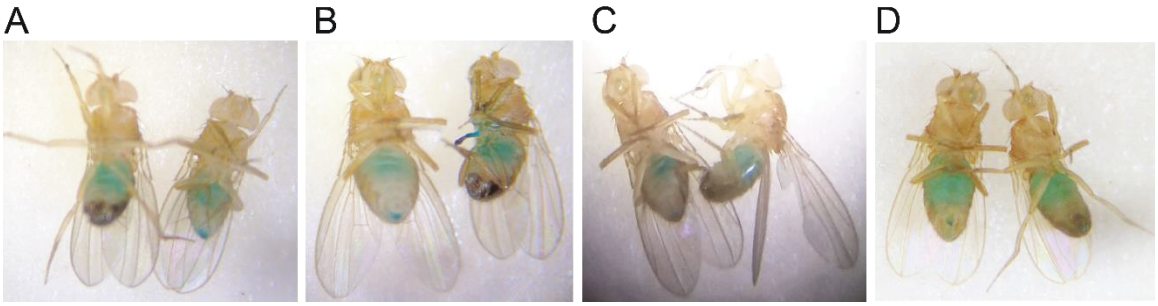

Figure E2. Representative image for *Drosophila* adults (male and female) maintained on apple juice containing halotyrosines and FD&C blue 1 dye for 5 days. (A) $CGG6279^{Del}$ strain (B) $CG6279^{E154Q}$ strain (C) $CG6279^{WT}$ strain (D) *yw* strain.

Table E2. *Drosophila* larval survival to adulthood (50 larvae per vial).

Compound Added	Strain	No of adults eclosed	Average	Standard Deviation
300 µM I-Tyr	*CG6279^{Del}*	0	0	0
		0		
		0		
	CG6279^{E154Q}	0	0	0
		0		
		0		
	CG6279^{WT}	0	1	1
		0		
		2		
		2		
	yw	4	9	5
		8		
		14		
200 µM I-Tyr	*CG6279^{Del}*	4	8	8
		18		
		0		
		8		
	CG6279^{E154Q}	0	0	0
		0		
		0		
	CG6279^{WT}	6	14	7
		17		
		20		
		12		
	yw	24	29	5
		36		
		26		
		29		

Table E2. Continued.

Compound Added	Strain	No of adults eclosed	Average	Standard Deviation
150 µM I-Tyr	*CG6279*^{Del}	47	48	1
		49		
		48		
		48		
	CG6279^{E154Q}	15	16	5
		9		
		20		
		18		
	CG6279^{WT}	50	49	1
		48		
		48		
		49		
	yw	46	47	2
		49		
		48		
		45		
100 µM I-Tyr	*CG6279*^{Del}	48	47	4
		49		
		47		
		45		
	CG6279^{E154Q}	14	32	12
		41		
		37		
		36		
	CG6279^{WT}	40	46	8
		46		
		48		
		49		
	yw	47	47	1
		47		
		48		

Table E2. Continued.

Compound Added	Strain	No of adults eclosed	Average	Standard Deviation
1500 µM Br-Tyr	*CG6279^{Del}*	20	29	8
		35		
		32		
	CG6279^{E154Q}	0	4	3
		6		
		6		
	CG6279^{WT}	2	10	9
		19		
		10		
	yw	18	18	1
		18		
		19		
1500 µM Cl-Tyr	*CG6279^{Del}*	50	46	3
		44		
		44		
	CG6279^{E154Q}	45	44	4
		39		
		47		
	CG6279^{WT}	48	47	1
		46		
		48		
	yw	20	22	3
		25		
		22		

Table E2. Continued.

Compound Added	Strain	No of adults eclosed	Average	Standard Deviation
1500 µM Tyr	*CG6279*^{Del}	43	46	3
		49		
		47		
	CG6279^{E154Q}	43	43	1
		44		
		42		
	CG6279^{WT}	49	49	2
		50		
		47		
	yw	47	48	1
		49		
		47		
300 µM Nal	*CG6279*^{Del}	44	44	2
		45		
		43		
	CG6279^{E154Q}	44	44	3
		42		
		47		
	CG6279^{WT}	50	47	4
		43		
		48		
	yw	47	45	2
		43		
		45		

Table E2. Continued.

Compound Added	Strain	No of adults eclosed	Average	Standard Deviation
1500 µM NaBr	*CG6279*^{Del}	43	47	3
		49		
		49		
	CG6279^{E154Q}	49	49	2
		47		
		50		
	CG6279^{WT}	50	49	2
		49		
		47		
	yw	46	47	2
		46		
		49		

Table E3. *Drosophila* adult survival (50 adults per vial).

Compound Added	Strain	No of adults eclosed	Average	Standard Deviation
2000 µM I-Tyr	*CG6279*^{Del}	44	44	1
		44		
		43		
	CG6279^{E154Q}	43	43	1
		44		
		43		
	CG6279^{WT}	47	46	2
		47		
		44		
		47		
	yw	41	43	3
		43		
		46		
2000 µM Br-Tyr	*CG6279*^{Del}	38	44	5
		47		
		47		
	CG6279^{E154Q}	46	46	1
		45		
		47		
	CG6279^{WT}	46	46	2
		47		
		45		
	yw	48	43	3
		40		
		46		
2000 µM Cl-Tyr	*CG6279*^{Del}	41	41	3
		38		
		43		
	CG6279^{E154Q}	46	46	1
		46		
		45		
	CG6279^{WT}	45	45	1
		46		
		44		
	yw	48	47	1
		46		
		47		

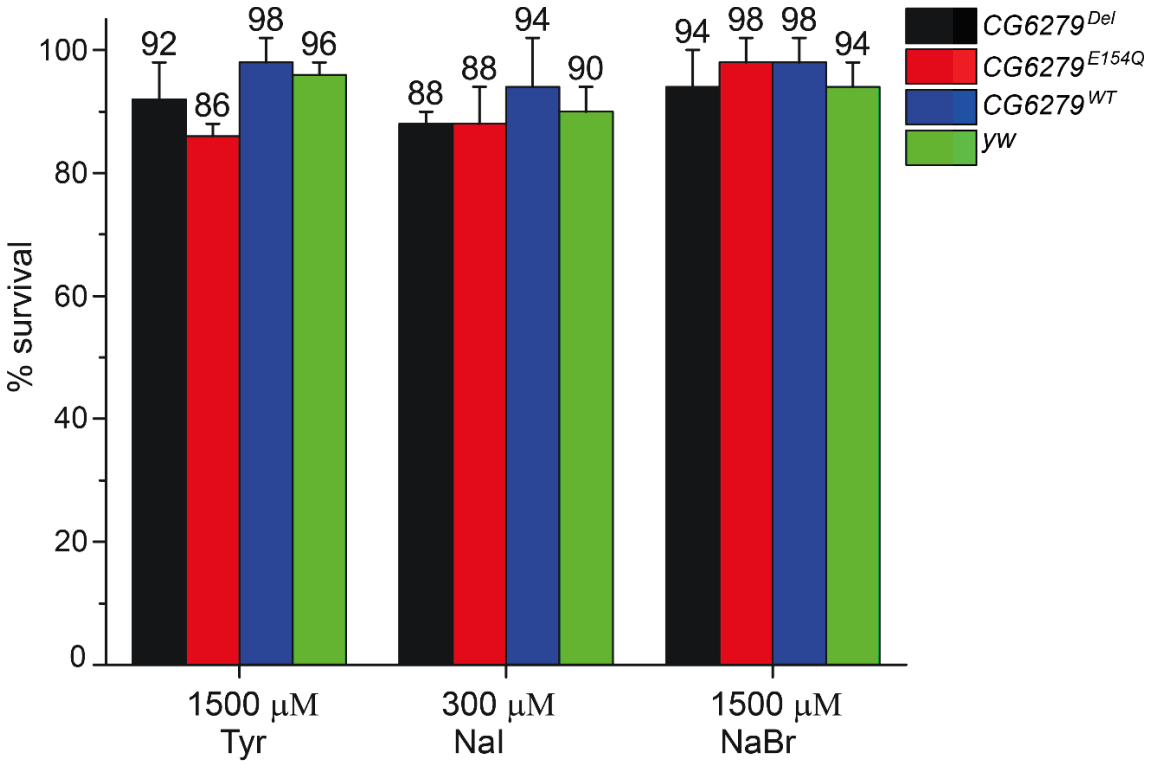

Figure E3. *Drosophila* larval survival rates when raised on media supplemented with appropriate control. Indicated numbers represent averages from 3 independent measurements and error bars represent their standard deviations. Bar plot is derived from data listed in Table E2.

Bibliography

1. Zimmermann, M. B. Iodine Deficiency. *Endocr. Rev.* 2009, 30, 376-408.

2. Zimmermann, M. B.; Jooste, P. L.; Pandav, C. S. Iodine-deficiency disorders. *Lancet* 2008, 372, 1251-1262.

3. Fuge, R. Sources of halogens in the environment, influences on human and animal health. *Environ. Geochem. Health* 1988, 10, 51-61.

4. Johnson, K. S.; Coale, K. H.; Jannasch, H. W. Analytical Chemistry in Oceanography. *Anal. Chem.* 1992, 64, 1065A-1075A.

5. Brent, G. A. Mechanisms of thyroid hormone action. *J. Clin. Invest.* 2012, 122, 3035-3043.

6. Miyata, K.; Ose, K. Thyroid Hormone-disrupting Effects and the Amphibian Metamorphosis Assay. *J. Toxicol. Pathol.* 2012, 25, 1-9.

7. Swingle, W. W. Iodine and Amphibian Metamorphosis. *Biological Bulletin* 1923, 45, 229-253.

8. Porazzi, P.; Calebiro, D.; Benato, F.; Tiso, N.; Persani, L. Thyroid gland development and function in the zebrafish model. *Mol. Cell. Endocrinol.* 2009, 312, 14-23.

9. Paris, M.; Brunet, F.; Markov, G.; Schubert, M.; Laudet, V. The amphioxus genome enlightens the evolution of the thyroid hormone signaling pathway. *Dev. Genes Evol.* 2008, 218, 667-680.

10. Paris, M.; Escriva, H.; Schubert, M.; Brunet, F.; Brtko, J.; Ciesielski, F.; Roecklin, D.; Vivat-Hannah, V.; Jamin, E. L.; Cravedi, J.-P.; Scanlan, T. S.; Renaud, J.-P.; Holland, N. D.; Laudet, V. Amphioxus postembryonic development reveals the homology of chordate metamorphosis. *Curr. Biol.* 2008, 18, 825-830.

11. Eskandari, S.; Loo, D. D. F.; Dai, G.; Levy, O.; Wright, E. M.; Carrasco, N. Thyroid Na^+/I^- Symporter: Mechanism, Stoichiometry, and Specificity. *J. Biol. Chem.* 1997, 272, 27230-27238.

12. Dohán, O.; De la Vieja, A.; Paroder, V.; Riedel, C.; Artani, M.; Reed, M.; Ginter, C. S.; Carrasco, N. The Sodium/Iodide Symporter (NIS): Characterization, Regulation, and Medical Significance. *Endocr. Rev.* 2003, 24, 48-77.

13. Rokita, S. E.; Adler, J. M.; McTamney, P. M.; Watson Jr, J. A. Efficient use and recycling of the micronutrient iodide in mammals. *Biochimie* 2010, 92, 1227-1235.

14. Gnidehou, S.; Caillou, B.; Talbot, M.; Ohayon, R.; Kaniewski, J.; Noël-Hudson, M. S.; Morand, S.; Agnangji, D.; Sezan, A.; Courtin, F.; Virion, A.; Dupuy, C. Iodotyrosine dehalogenase 1 (DEHAL1) is a transmembrane protein involved in the recycling of iodide close to the thyroglobulin iodination site. *FASEB J.* 2004, 18, 1574-1576.

15. Moreno, J. C.; Klootwijk, W.; van Toor, H.; Pinto, G.; D'Alessandro, M.; Lèger, A.; Goudie, D.; Polak, M.; Grüters, A.; Visser, T. J. Mutations in the Iodotyrosine Deiodinase Gene and Hypothyroidism. *N. Engl. J. Med.* 2008, 358, 1811-1818.

16. Roche, J.; Michel, R.; Michel, O.; Lissitzky, S. Sur la déshalogénation enzymatique des iodotyrosines par le corps thyroïde et sur son rôle physiologique. *Biochim. Biophys. Acta* 1952, 9, 161-169.

17. Goswami, A.; Rosenberg, I. N. Studies on a soluble thyroid Iodotyrosine deiodinase: Activation by NADPH and electron carriers. *Endocrinology* 1977, 101, 331-341.

18. Rosenberg, I. N.; Goswami, A. Purification and characterization of a flavoprotein from bovine thyroid with iodotyrosine deiodinase activity. *J. Biol. Chem.* 1979, 254, 12318-12325.

19. McTamney, P. M.; Rokita, S. E. A mammalian reductive deiodinase has broad power to dehalogenate chlorinated and brominated substrates. *J. Am. Chem. Soc.* 2009, 131, 14212-14213.

20. Mohn, W. W., Tiedje, J. M. Microbial reductive dehalogenation. *Microbiol. Rev.* 1992, 56, 482-507.

21. St. Germain, D. L.; Galton, V. A.; Hernandez, A. Defining the Roles of the Iodothyronine Deiodinases: Current Concepts and Challenges. *Endocrinology* 2009, 150, 1097-1107.

22. Berry, M. J.; Kieffer, J. D.; Harney, J. W.; Larsen, P. R. Selenocysteine confers the biochemical properties characteristic of the type I iodothyronine deiodinase. *J. Biol. Chem.* 1991, 266, 14155-14158.

23. Friedman, J. E.; Watson, J. A.; Lam, D. W. H.; Rokita, S. E. Iodotyrosine deiodinase is the first mammalian member of the NADH oxidase/flavin reductase superfamily. *J. Biol. Chem.* 2006, 281, 2812-2819.

24. Roldán, M. D.; Pérez-Reinado, E.; Castillo, F.; Moreno-Vivián, C. Reduction of polynitroaromatic compounds: the bacterial nitroreductases. *FEMS Microbiol. Rev.* 2008, 32, 474-500.

25. Watson, J. A.; McTamney, P. M.; Adler, J. M.; Rokita, S. E. Flavoprotein Iodotyrosine Deiodinase Functions without Cysteine Residues. *ChemBioChem* 2008, 9, 504-506.

26. Watson, J. A. Insight into the Structure and Mechanism of Iodotyrosine Deiodinase, the First Mammalian Member of the NADH Oxidase / Flavin Reductase Superfamily. Ph.D. Dissertation, University of Maryland, College Park, 2006.

27. McTamney, P. M. Catalytic features of the iodine salvaging enzyme iodotyrosine deiodinase. Ph.D. Dissertation, University of Maryland, College Park, 2009.

28. Thomas, S. R.; McTamney, P. M.; Adler, J. M.; LaRonde-LeBlanc, N.; Rokita, S. E. Crystal structure of iodotyrosine deiodinase, a novel flavoprotein responsible for iodide salvage in thyroid glands. *J. Biol. Chem.* 2009, 284, 19659-19667.

29. Buss, J. M.; McTamney, P. M.; Rokita, S. E. Expression of a soluble form of iodotyrosine deiodinase for active site characterization by engineering the native membrane protein from *Mus musculus*. *Protein Sci.* 2012, 21, 351-361.

30. Buss, J. M. Iodotyrosine deiodinase from selected phyla engineered for bacterial expression. Ph.D. Dissertation, University of Maryland, College Park, 2012.

31. Hu, J.; Chuenchor, W.; Rokita, S. E. A Switch between One- and Two-electron Chemistry of the Human Flavoprotein Iodotyrosine Deiodinase Is Controlled by Substrate. *J. Biol. Chem.* 2015, 290, 590-600.

32. Mossessova, E.; Lima, C. D. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol. Cell* 2000, 5, 865-876.

33. Taga, M. E.; Larsen, N. A.; Howard-Jones, A. R.; Walsh, C. T.; Walker, G. C. BluB cannibalizes flavin to form the lower ligand of vitamin B_{12}. *Nature* 2007, 446, 449-453.

34. Hecht, H. J.; Erdmann, H.; Park, H. J.; Sprinzl, M.; Schmid, R. D. Crystal structure of NADH oxidase from Thermus thermophilus. *Nat. Struct. Biol.* 1995, 2, 1109-1114.

35. Tanner, J. J.; Lei, B.; Tu, S. C.; Krause, K. L. Flavin reductase P: Structure of a dimeric enzyme that reduces flavin. *Biochemistry* 1996, 35, 13531-13539.

36. Rosenberg, I. N.; Goswami, A. Iodotyrosine deiodinase from bovine thyroid. *Methods Enzymol.* 1984, 107, 488-500.

37. Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997, 25, 3389-3402.

38. Crockford, S. J. Evolutionary roots of iodine and thyroid hormones in cell-cell signaling. *Integr. Comp. Biol.* 2009, 49, 155-166.

39. Glasner, M. E.; Gerlt, J. A.; Babbitt, P. C. Evolution of enzyme superfamilies. *Curr. Opin. Chem. Biol.* 2006, 10, 492.

40. Babbitt, P. C.; Gerlt, J. A. Understanding Enzyme Superfamilies. *J. Biol. Chem.* 1997, 272, 30591-30594.

41. Jones, C. E.; Brown, A. L.; Baumann, U. Estimating the annotation error rate of curated GO database sequence annotations. *BMC Bioinfomatics* 2007, 8, 170.

42. Schnoes, A. M.; Brown, S. D.; Dodevski, I.; Babbitt, P. C. Annotation Error in Public Databases: Misannotation of Molecular Function in Enzyme Superfamilies. *PLoS Comput. Biol.* 2009, 5, e1000605.

43. Gasteiger E.; Hoogland C.; Gattiker A.; Duvaud S.; Wilkins M.R.; Appel R.D.; Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server, in *The Proteomics Protocols Handbook* (Walker J. M., Ed.), 2006, pp 571-607, Humana Press, NY.

44. Jacek, K. Fluorometric analyses of riboflavin and its coenzymes. *Methods Enzymol.* 1971, 18, 253-285.

45. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004, 32, 1792-1797.

46. Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *J. Mol. Biol.* 2001, 305, 567-580.

47. Yu, T. Y.; Mok, K. C.; Kennedy, K. J.; Valton, J.; Anderson, K. S.; Walker, G. C.; Taga, M. E. Active site residues critical for flavin binding and 5,6-dimethylbenzimidazole biosynthesis in the flavin destructase enzyme BluB. *Protein Sci.* 2012, 21, 839-849.

48. Sayers, E. W.; Barrett, T., et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 2009, 37, D5-15.

49. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 2003, 13, 2498-2504.

50. Atkinson, H. J.; Morris, J. H.; Ferrin, T. E.; Babbitt, P. C. Using Sequence Similarity Networks for Visualization of Relationships Across Diverse Protein Superfamilies. *PLoS One* 2009, 4, e4345.

51. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 2010, 59, 307-321.

52. Chevenet, F.; Brun, C.; Banuls, A.-L.; Jacq, B.; Christen, R. TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinfomatics* 2006, 7, 439.

53. Eales, J. G. Iodine Metabolism and Thyroid-Related Functions in Organisms Lacking Thyroid Follicles: Are Thyroid Hormones also Vitamins? *Proc. Soc. Exp. Biol. Med.* 1997, 214, 302-317.

54. Heyland, A.; Moroz, L. L. Cross-kingdom hormonal signaling: an insight from thyroid hormone functions in marine larvae. *J. Exp. Biol.* 2005, 208, 4355-4361.

55. Thyagaraja, B. S.; Masler, E. P.; Kelly, T. J.; Borkovec, A. B. Thyroxine-induced changes in ovarian protein and ecdysteroid levels in the silkworm, Bombyx mori L.: Effect on ovarian maturation and egg production. *Comp. Biochem. Physiol.* 1993, 104A, 247-253.

56. Davey, K. G. Do thyroid hormones function in insects? *Insect Biochem. Mol. Biol.* 2000, 30, 877-884.

57. Flatt, T.; Moroz, L. L.; Tatar, M.; Heyland, A. Comparing Thyroid and Insect Hormone Signaling. *Integr. Comp. Biol.* 2006, 46, 777-794.

58. Silverstone, M.; Galton, V. A.; Ingbar, S. H. Observations concerning the metabolism of iodine by polyps of Aurelia aurita. *Gen. Comp. Endocrinol.* 1978, 34, 132-140.

59. Amachi, S.; Mishima, Y.; Shinoyama, H.; Muramatsu, Y.; Fujii, T. Active Transport and Accumulation of Iodide by Newly Isolated Marine Bacteria. *Appl. Environ. Microbiol.* 2005, 71, 741-745.

60. Li, H.-P.; Brinkmeyer, R.; Jones, W. L.; Zhang, S.; Xu, C.; Schwehr, K. A.; Santschi, P. H.; Kaplan, D. I.; Yeager, C. M. Iodide Accumulation by Aerobic Bacteria Isolated from Subsurface Sediments of a 129I-Contaminated Aquifer at the Savannah River Site, South Carolina. *Appl. Environ. Microbiol.* 2011, 77, 2153-2160.

61. Macheroux, P.; Kappes, B.; Ealick, S. E. Flavogenomics – a genomic and structural view of flavin-dependent proteins. *FEBS J.* 2011, 278, 2625-2634.

62. Bobyk, K. D.; Ballou, D. P.; Rokita, S. E. Rapid Kinetics of Dehalogenation Promoted by Iodotyrosine Deiodinase from Human Thyroid. *Biochemistry* 2015, 54, 4487-4494.

63. Hunt, S. Halogenated tyrosine derivatives in invertebrate scleroproteins: Isolation and identification. *Methods Enzymol.* 1984, 107, 413-438.

64. Hopkins, T. L.; Kramer, K. J. Insect Cuticle Sclerotization. *Annu. Rev. Entomol.* 1992, 37, 273-302.

65. Attrill, H.; Falls, K.; Goodman, J. L.; Millburn, G. H.; Antonazzo, G.; Rey, A. J.; Marygold, S. J. FlyBase: establishing a Gene Group resource for *Drosophila melanogaster*. *Nucleic Acids Res.* 2016, 44, D786-D792.

66. Warner, J. R.; Copley, S. D. Pre-Steady-State Kinetic Studies of the Reductive Dehalogenation Catalyzed by Tetrachlorohydroquinone Dehalogenase. *Biochemistry* 2007, 46, 13211-13222.

67. McMillen, D. F.; Golden, D. M. Hydrocarbon Bond Dissociation Energies. *Ann. Rev. Phys. Chem.* 1982, 33, 493-532.

68. Odar, C.; Winkler, M.; Wiltschi, B. Fluoro amino acids: A rarity in nature, yet a prospect for protein engineering. *Biotechnol. J.* 2015, 10, 427-446.

69. Wheeler, B. M. Halogen metabolism of Drosophila gibberosa. I. Iodine metabolism studied by means of I^{131}. *J. Exp. Zool.* 1950, 115, 83-107.

70. Mani, A. R.; Moreno, J. C.; Visser, T. J.; Moore, K. P. The metabolism and de-bromination of bromotyrosine in vivo. *Free Radic. Biol. Med.* 2016, 90, 243-251.

71. McCall Scott, A.; Cummings, C. F.; Bhave, G.; Vanacore, R.; Page-McCaw, A.; Hudson, B. G. Bromine Is an Essential Trace Element for Assembly of Collagen IV Scaffolds in Tissue Development and Architecture. *Cell* 2014, 157, 1380-1392.

72. Andersen, S. O. Chlorinated tyrosine derivatives in insect cuticle. *Insect Biochem. Mol. Biol.* 2004, 34, 1079-1087.

73. Mukherjee, A.; Rokita, S. E. Single Amino Acid Switch between a Flavin-Dependent Dehalogenase and Nitroreductase. *J. Am. Chem. Soc.* 2015, 137, 15342-15345.

74. Ralston, I. M.; Dunford, H. B. Horseradish peroxidase. XLII. Oxidations of L-tyrosine and 3,5-diiodo-L-tyrosine by compound II. *Can. J. Biochem.* 1980, 58, 1270-1276.

75. Das, T. N. Redox Chemistry of 3-Iodotyrosine in Aqueous Medium. *J. Phys. Chem. A* 1998, 102, 426-433.

76. Jennings, B. H. Drosophila: a versatile model in biology & medicine. *Materials Today* 2011, 14, 190-195.

77. Ashburner, M.; Bergman, C. M. Drosophila melanogaster: A case study of a model genomic sequence and its consequences. *Genome Res.* 2005, 15, 1661-1667.

78. St Johnston, D. The art and design of genetic screens: Drosophila melanogaster. *Nat. Rev. Genet.* 2002, 3, 176-188.

79. Adams, M. D.; Celniker, S. E., et al. The Genome Sequence of Drosophila melanogaster. *Science* 2000, 287, 2185-2195.

80. Adams, M. D.; Sekelsky, J. J. From sequence to phenotype: reverse genetics in Drosophila melanogaster. *Nat. Rev. Genet.* 2002, 3, 189-198.

81. Ashburner M.; Golic K.G.; Hawley R. S. Drosophila: A Laboratory Handbook, 2nd Ed. 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

82. Kirkpatrick, M. How and Why Chromosome Inversions Evolve. *PLoS Biol.* 2010, 8, e1000501.

83. Rubin, G. M.; Lewis, E. B. A Brief History of Drosophila's Contributions to Genome Research. *Science* 2000, 287, 2216-2218.

84. Boutros, M.; Ahringer, J. The art and design of genetic screens: RNA interference. *Nat. Rev. Genet.* 2008, 9, 554-566.

85. Brand, A. H.; Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 1993, 118, 401-415.

86. Duffy, J. B. GAL4 system in Drosophila: A fly geneticist's swiss army knife. *Genesis* 2002, 34, 1-15.

87. Chintapalli, V. R.; Wang, J.; Dow, J. A. T. Using FlyAtlas to identify better Drosophila melanogaster models of human disease. *Nat. Genet.* 2007, 39, 715-720.

88. Roy, S.; Ernst, J., et al. Identification of Functional Elements and Regulatory Circuits by Drosophila modENCODE. *Science* 2010, 330, 1787-1797.

89. Eun, S. H.; Stoiber, P. M.; Wright, H. J.; McMurdie, K. E.; Choi, C. H.; Gan, Q.; Lim, C.; Chen, X. MicroRNAs downregulate Bag of marbles to ensure proper terminal differentiation in the Drosophila male germline. *Development* 2013, 140, 23-30.

90. White-Cooper, H. Molecular mechanisms of gene regulation during Drosophila spermatogenesis. *Reproduction* 2010, 139, 11-21.

91. Kiger, A. A.; White-Cooper, H.; Fuller, M. T. Somatic support cells restrict germline stem cell self-renewal and promote differentiation. *Nature* 2000, 407, 750-754.

92. Schulz, C. In Situ Hybridization to Drosophila Testes. *CSH Protoc.* 2007, doi:10.1101/pdb.prot4764.

93. Morris, C. A.; Benson, E.; White-Cooper, H. Determination of gene expression patterns using in situ hybridization to Drosophila testes. *Nat. Protoc.* 2009, 4, 1807-1819.

94. Vert, J.-P.; Foveau, N.; Lajaunie, C.; Vandenbrouck, Y. An accurate and interpretable model for siRNA efficacy prediction. *BMC Bioinformatics* 2006, 7, 520.

95. Kulkarni, M. M.; Booker, M.; Silver, S. J.; Friedman, A.; Hong, P.; Perrimon, N.; Mathey-Prevot, B. Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. *Nat. Methods* 2006, 3, 833-838.

96. Ni, J.-Q.; Perrimon, N. Cloning hairpins into VALIUM20 and VALIUM22. *Protocols at the Transgenic RNAi Project (TRIP)*, http://www.flyrnai.org/supplement/2ndGenProtocol.pdf.

97. Ayyar, S.; Jiang, J.; Collu, A.; White-Cooper, H.; White, R. A. H. Drosophila TGIF is essential for developmentally regulated transcription in spermatogenesis. *Development* 2003, 130, 2841-2852.

98. Moon, S.; Cho, B.; Min, S.-H.; Lee, D.; Chung, Y. D. The THO complex is required for nucleolar integrity in Drosophila spermatocytes. *Development* 2011, 138, 3835-3845.

99. Dietzl, G.; Chen, D.; Schnorrer, F.; Su, K.-C.; Barinova, Y.; Fellner, M.; Gasser, B.; Kinsey, K.; Oppel, S.; Scheiblauer, S.; Couto, A.; Marra, V.; Keleman, K.; Dickson, B. J. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature* 2007, 448, 151-156.

100. Haley, B.; Hendrix, D.; Trang, V.; Levine, M. A simplified miRNA-based gene silencing method for Drosophila melanogaster. *Dev. Biol.* 2008, 321, 482-490.

101. Bassett, A. R.; Liu, J.-L. CRISPR/Cas9 and Genome Editing in Drosophila. *J. Genet. Genomics* 2014, 41, 7-19.

102. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D. A.; Horvath, P. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* 2007, 315, 1709-1711.

103. Sander, J. D.; Joung, J. K. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 2014, 32, 347-355.

104. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J. A.; Charpentier, E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 2012, 337, 816-821.

105. Urnov, F. D.; Rebar, E. J.; Holmes, M. C.; Zhang, H. S.; Gregory, P. D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* 2010, 11, 636-646.

106. Joung, J. K.; Sander, J. D. TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell. Biol.* 2013, 14, 49-55.

107. Doudna, J. A.; Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* 2014, 346, 1258096.

108. Ren, X.; Sun, J.; Housden, B. E.; Hu, Y.; Roesel, C.; Lin, S.; Liu, L.-P.; Yang, Z.; Mao, D.; Sun, L.; Wu, Q.; Ji, J.-Y.; Xi, J.; Mohr, S. E.; Xu, J.; Perrimon, N.; Ni, J.-Q. Optimized gene editing technology for Drosophila melanogaster using germ line-specific Cas9. *Proc. Natl. Acad. Sci.* 2013, 110, 19012-19017.

109. Bassett, A. R.; Tibbit, C.; Ponting, C. P.; Liu, J. L. Highly Efficient Targeted Mutagenesis of Drosophila with the CRISPR/Cas9 System. *Cell Rep.* 2013, 4, 220-228.

110. Gratz, S. J.; Cummings, A. M.; Nguyen, J. N.; Hamm, D. C.; Donohue, L. K.; Harrison, M. M.; Wildonger, J.; O'Connor-Giles, K. M. Genome Engineering of Drosophila with the CRISPR RNA-Guided Cas9 Nuclease. *Genetics* 2013, 194, 1029-1035.

111. Yu, Z.; Ren, M.; Wang, Z.; Zhang, B.; Rong, Y. S.; Jiao, R.; Gao, G. Highly Efficient Genome Modifications Mediated by CRISPR/Cas9 in Drosophila. *Genetics* 2013, 195, 289-291.

112. Gratz, S. J.; Ukken, F. P.; Rubinstein, C. D.; Thiede, G.; Donohue, L. K.; Cummings, A. M.; O'Connor-Giles, K. M. Highly Specific and Efficient CRISPR/Cas9-Catalyzed Homology-Directed Repair in Drosophila. *Genetics* 2014, 196, 961-971.

113. Port, F.; Chen, H.-M.; Lee, T.; Bullock, S. L. Optimized CRISPR/Cas tools for efficient germline and somatic genome engineering in Drosophila. *Proc. Natl. Acad. Sci.* 2014, 111, E2967-E2976.

114. Port, F. Cloning gRNA expression vectors with pCFD3. http://www.crisprflydesign.org/.

115. Bloomington Drosophila Stock Center: Deficiency Kit information. [http://flystocks.bio.indiana.edu/Browse/df/dfkit-info.htm].

116. Xie, J.; Wooten, M.; Tran, V.; Chen, B.-C.; Pozmanter, C.; Simbolon, C.; Betzig, E.; Chen, X. Histone H3 Threonine Phosphorylation Regulates Asymmetric Histone Inheritance in the Drosophila Male Germline. *Cell* 2015, 163, 920-933.

117. Cook, R. K.; Christensen, S. J.; Deal, J. A.; Coburn, R. A.; Deal, M. E.; Gresens, J. M.; Kaufman, T. C.; Cook, K. R. The generation of chromosomal deletions to provide extensive coverage and subdivision of the Drosophila melanogaster genome. *Genome Biol.* 2012, 13, R21.

118.Valverde-R, C.; Orozco, A.; Becerra, A.; Jeziorski, M. C.; Villalobos, P.; Solis-S, J. C. Halometabolites and Cellular Dehalogenase Systems: An Evolutionary Perspective. *Int. Rev. Cytol.* 2004, 234, 143-199.

119. Fitzpatrick, P. F. The pH dependence of binding of inhibitors to bovine adrenal tyrosine hydroxylase. *J. Biol. Chem.* 1988, 263, 16058-16062.

120. Udenfriend, S.; Zaltzman-Nirenberg, P.; Nagatsu, T. Inhibitors of purified beef adrenal tyrosine hydroxylase. *Biochem. Pharmacol.* 1965, 14, 837-845.

121. Fitzpatrick, P. F. Studies of the rate-limiting step in the tyrosine hydroxylase reaction: alternate substrates, solvent isotope effects, and transition-state analogs. *Biochemistry* 1991, 30, 6386-6391.

122. Yamamoto, S.; Seto, E. S. Dopamine Dynamics and Signaling in *Drosophila*: An Overview of Genes, Drugs and Behavioral Paradigms. *Exp. Anim.* 2014, 63, 107-119.

123. Yellman, C.; Tao, H.; He, B.; Hirsh, J. Conserved and sexually dimorphic behavioral responses to biogenic amines in decapitated Drosophila. *Proc. Natl. Acad. Sci.* 1997, 94, 4131-4136.

124. Ueno, T.; Tomita, J.; Tanimoto, H.; Endo, K.; Ito, K.; Kume, S.; Kume, K. Identification of a dopamine pathway that regulates sleep and arousal in Drosophila. *Nat. Neurosci.* 2012, 15, 1516-1523.

125. Selcho, M.; Pauls, D.; Han, K.-A.; Stocker, R. F.; Thum, A. S. The Role of Dopamine in Drosophila Larval Classical Olfactory Conditioning. *PLoS ONE* 2009, 4, e5897.

126. Neckameyer, W. S. Dopamine Modulates Female Sexual Receptivity in Drosophila Melanogaster. *J. Neurogenet.* 1998, 12, 101-114.

127. Xiao, N.; Venton, B. J. Characterization of dopamine releasable and reserve pools in *Drosophila* larvae using ATP/P2X2-mediated stimulation. *J. Neurochem.* 2015, 134, 445-454.

128. Neckameyer, W. S. Multiple Roles for Dopamine in Drosophila Development. *Dev. Biol.* 1996, 176, 209-219.

129. Ashburner M.; Roote J. Laboratory culture of Drosophila, in *Drosophila Protocols* (Sullivan W.; Ahburner M; Hawley S. R., Eds.) 2000, pp 588-599, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

130. Sinden, D.; Badgett, M.; Fry, J.; Jones, T.; Palmen, R.; Sheng, X.; Simmons, A.; Matunis, E.; Wawersik, M. Jak-STAT regulation of cyst stem cell development in the Drosophila testis. *Dev. Biol.* 2012, 372, 5-16.

131. Law, J. H.; Wells, M. A. Insects as biochemical models. *J. Biol. Chem.* 1989, 264, 16335-16338.

132. Arrese, E. L.; Soulages, J. L. Insect Fat Body: Energy, Metabolism, and Regulation. *Annu. Rev. Entomol.* 2009, 55, 207-225.

133. O'Connor-Giles, K., Personal communication to FlyBase, FBrf0224309. In 2014.

134. Eleftherianos, I.; More, K.; Spivack, S.; Paulin, E.; Khojandi, A.; Shukla, S. Nitric Oxide Levels Regulate the Immune Response of Drosophila melanogaster Reference Laboratory Strains to Bacterial Infections. *Infect. Immun.* 2014, 82, 4169-4181.

135. Ryuda, M.; Tsuzuki, S.; Tanimura, T.; Tojo, S.; Hayakawa, Y. A gene involved in the food preferences of larval Drosophila melanogaster. *J. Insect Physiol.* 2008, 54, 1440-1445.

Curriculum Vitae

Abhishek V Phatarphekar
Johns Hopkins University
Dept. of Chemistry, Remsen 122
Baltimore, MD 21218
abhishekph@gmail.com

Education

Johns Hopkins University, School of Arts and Sciences, Baltimore, MD
PhD Candidate (Chemistry) – expected graduation August, 2016

University of Pennsylvania, SEAS, Philadelphia, PA
Master of Biotechnology, December 2008
Cumulative GPA: 3.92 on a scale of 4.0

Mumbai University Institute of Chemical Technology, Mumbai, India
Bachelor of Pharmaceutical Sciences, July 2007
Grade I with Distinction

Research and Work Experience

Rokita Lab, Johns Hopkins University, Jul 2012 - present, University of Maryland, College Park, Aug 2011 - Jul 2012
PhD research involves catalytic characterization of halotyrosine dehalogenase orthologs from eukaryotes to prokaryotes. Cloning, expression, purification, site directed mutagenesis and characterization of catalysis by the *Drosophila* halotyrosine dehalogenase. In vivo functional analysis of the *Drosophila* halotyrosine dehalogenase through genome editing using CRISPR in collaboration with Dr Xin Chen at the Biology Dept., Johns Hopkins University.

Research Associate III, ImmunoGen, Inc., Waltham, MA, Jan 2011 – May 2011
Responsibilites: Optimization of antibody-drug conjugation techniques using ImmunoGen's proprietary linker technology, antibody-drug conjugate (ADC) purification, physico-chemical characterization of ADCs using analytical tools such as HPLC and spectroscopy and formulation of ADCs for animal testing.

Staff Biologist, GlycoFi (subsidiary of Merck & Co., Inc), Lebanon NH, Apr 2009 – Dec 2010
Responsibilites: Develop protocols for purification of recombinant proteins including monoclonal antibodies expressed in glycoengineered yeast *Pichia Pastoris* using AKTA FPLC systems (developed multistep purification processes for proteins of interest). Delivered protocols to Process Development group for scale up. Development and execution of ELISA procedures for screening to determine protein expression.

Summer Intern, Genzyme Corporation, Waltham, MA, Jun 2008 – Aug 2008
Successfully completed the development of a highly sensitive LC-MS/MS method to accurately quantify 1,25 $(OH)_2$ Vitamin D_2 in rat serum at the pg/ml scale.

Research Assistant, University of Pennsylvania, **Jan2008 - June 2008**
High Throughput Screening of compound library for identifying inhibitors of fibrin polymerization (1536 and 384 well plate format). Screening was performed using an automated HTS platform and robotic liquid handlers.

Publications

Abhishek Phatarphekar and Steven E. Rokita (2016) Functional analysis of iodotyrosine deiodinase from *Drosophila melanogaster*. Submitted

Abhishek Phatarphekar, Jennifer M. Buss and Steven E. Rokita. (2014) Iodotyrosine deiodinase: a unique flavoprotein present in organisms of diverse phyla. *Mol. Biosyst.* 10, 86-92.
Highlights: "I branches out", Nat. Chem. Biol, 2014, 10, Pg 2

Sirimas Sudsakorn, Abhishek Phatarphekar, Thomas O'Shea and Hanlan Liu. (2011) Determination of 1,25 dihydroxyvitamin D_2 in rat serum using liquid chromatography with tandem mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 879, 139-45.

Honors and Awards

Harry and Cleio Greer Fellowship, Johns Hopkins University - Chemistry Department, June 2014 – May 2015

Dean's Fellowship, University of Maryland, August 2011-June 2012

Merck Special Achievement Award in the "Innovative Technologies/Process Improvements" category – February 2010